READING SKILL AND COMPONENTS OF WORD KNOWLEDGE AFFECT EYE MOVEMENTS DURING READING

by

Jessica Nelson

B.A., Rice University, 2001

M.S., University of Pittsburgh, 2005

Submitted to the Graduate Faculty of

the Department of Psychology in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

DEPARTMENT OF PSYCHOLOGY

This dissertation was presented

by

Jessica Nelson

It was defended on

August 9th, 2010

and approved by

Charles A. Perfetti, Department of Psychology

Tessa Warren, Department of Psychology

Natasha Tokowicz, Department of Psychology

Isabel Beck, School of Education

Dissertation Director: Charles A. Perfetti, Department of Psychology

READING SKILL AND COMPONENTS OF WORD KNOWLEDGE AFFECT EYE MOVEMENTS DURING READING

Jessica Nelson, PhD

University of Pittsburgh, 2010

Four studies were conducted to examine the effects of individual skill differences among adult readers, as well as the effects of the quality of their lexical knowledge, on eye movements during reading. Study 1 defined dimensions of adult reader variability via a factor analysis of a database of adult reading assessments. Five dimensions emerged from the analysis reflecting: (1) speed/reading experience (expertise), (2) sublexical skills, (3) accuracy, (4) learning/memory, and (5) amount of "casual reading". Study 2 examined the effects of each of the dimensions of variability on eye movements during paragraph reading. Expert readers read words more quickly, especially less frequent words. Readers with good sublexical skills exhibited faster reading on early fixations, especially for more frequent words. These results suggest that individual differences in reading may be largely based on differences in the quality of lexical representations, with experienced readers having more knowledge of low frequency words, and readers with good sublexical skills having more unitized representations of frequent words. Study 3 employed a training paradigm in order to control the quality of lexical knowledge readers had along orthographic, phonological, and meaning dimensions. Orthographic and phonological training affected first pass reading measures, and phonological and meaning training affected second pass measures. The direction and strength of the training effects were mediated by individual differences between readers in their ability to learn from experiences. Study 4 examined the effects of components of word knowledge on eye movements by testing readers' orthographic, phonological, and meaning knowledge of words they read in context. Results confirmed the results of Study 3, with

orthographic and phonological knowledge affecting early fixations, and meaning knowledge affecting re-reading. Studies 3 and 4 also showed a pattern of faster first fixations with more rereading for the words with the list familiar forms. Results from the set of experiments show that each component of word knowledge affects eye movements uniquely, and that individual differences between adult readers, resulting in differences in the quality of lexical representations these readers have, account for variability in patterns of reading above and beyond attributes of the text being read.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I am so pleased to have the chance to thank all of the important people in my life who have helped me along the path to completing my dissertation in the field of reading research.

My mother, for sparking my interest in reading at such a young age. I still remember our homemade books about eating at Duff's and the way we cut out and captioned my favorite magazine photos.

My father, for always encouraging sound arguments and rigorous, logical thinking. I have no doubt that this made me into the scientific thinker I am today.

David, for taking me to stores with cool science toys and gadgets, and encouraging my interest in all things science.

Lucy, for being someone who could always relate what I was working on to the "real world." Your experience with reading education has helped me understand the importance of research in this field!

Julie, Emily, and Caroline, whose pride in me I value immensely, and who I know will always want the best for me. Thank you for supporting me always!

Reed Taylor, for all of the encouragement, for picking up all of the slack while I worked, for putting up with me with a kind heart and infinite understanding, for helping me cope, and for being so lovable!

The amazing teachers I have been so lucky to have: Merrie Southgate, Denie Ravenel, and Margaret Jollow, who made science my favorite subject. Dana Mobley and Jackie Joye who instilled in me a love of words, and who probably don't know that I wrote a college entrance essay claiming that my favorite book was the thesaurus, from the Latin word for "treasure". Jeff Wragg and "Mama Vincent" who made a career in research seem so cool. Spike Gildea, who taught linguistics with such passion that I changed my major. And Ian

Barland, who taught me that it is possible to learn the names of and give personal attention to everyone in a 100 person class – I will strive to be this kind of teacher.

My irreplaceable friends, Ryan and Laura, who have been so supportive of me throughout this process.

Michelle, my partner in crime, for being the kind of friend who is impossible not to have fun with.

All of the friends I have made in graduate school, who have kept me sane and have always been willing to help me out, especially Michal, Liz, Dimitrios, Eebie, Sara, D.J., J.C., and Tamar.

Marianne Vu for the unbelievable amount of help she provided with the dissertation experiments, and her cheerful loyalty to the project – I definitely couldn't have done it without her. And Caitlyn Hobaugh, who voluntarily spent beautiful summer days running participants for this project.

Of course, the graduate faculty who have inspired me, coached me, taught me to be a scientist, served on my committees, and have been my friends: Julie Fiez, Natasha Tokowicz, Isabel Beck, Tessa Warren, and Erik Reichle.

And most of all, my advisor, Charles Perfetti. I have learned so much from him, including the meanings of a plethora of rare words. His science, wisdom, and accomplishments are inspirational, and I feel fortunate to have been trained by the best. I must also note that I have fully enjoyed both his wry sense of humor and the way he let me push that sense of humor to the limits with my antics. Thank you, Chuck, for being both an advisor and a friend.

1.0 INTRODUCTION

1.1 OVERVIEW

This introduction provides background illustrating the following claims: (1) Two important components of skilled reading are the ability to identify and understand words (a) accurately and (b) rapidly and automatically. (2) These abilities develop from (a) a reader's previous experience with the particular word being read and (b) the reader's ability to learn from experience with words. (3) Both experience and ability to learn from experience are important sources of individual differences. This background motivates the research goal of the dissertation: to explore the effects of individual differences among normal adult readers (in terms of skills, experience, and lexical quality) on fluent silent reading at the word level.

The following research goals are undertaken in the thesis: (1) Identify major dimensions along which reading skill and experience vary among normal adult readers using a data-driven approach (Study 1), (2) understand whether and how those individual differences help predict patterns of on-line reading (Studies 2 & 3) , and (3) more directly measure the consequences of the quality of lexical representations for reading behavior (Study 4).

1.2 BACKGROUND

1.2.1 Lexical Quality and Comprehension

The task of reading can be broken down into a myriad of interdependent and temporally coordinated cognitive processes. Perhaps the most basic of these processes is understanding the words that constitute the text. Although higher level skills such as inference making and text integration are also required for skilled comprehension, there are few readers with good comprehension skill despite poor word-level skills. Thus, it has been hypothesized that a reader's vocabulary knowledge and the quality of a reader's lexical representations contributes to the reader's comprehension ability. This idea is known as the Lexical Quality Hypothesis (Perfetti & Hart, 2002, 2001; Perfetti, 2007b).

Correlational data show that, indeed, readers with high comprehension ability also tend to have high levels of word knowledge and word reading skills. Based on overlapping databases of reader scores on a battery of reading tests, both Hart (2005) and Landi (2005) found a strong relationship between scores on tests of lexical knowledge and tests of comprehension. Hart (2005) found that 65% of her sample was either above or below the median for both decoding/spelling scores and comprehension/vocabulary scores. In other words, the majority of readers have test scores that are consistent between the word form level and the word and text comprehension level. Additionally, Landi (2005) found that only 9% of her sample scored in the top 50% for comprehension but the bottom 50% on a lexical form factor. This means that few people are good comprehenders with poor word-level knowledge and skills.

Both Perfetti and Hart (2002) and Landi (2005) performed principal component analyses (PCA's) on the test scores to determine what functionally distinct skills or knowledge the various tests measured. Both studies determined that two factors, at least for skilled readers, accounted for a majority of the variance. One factor was a form factor related to spelling and decoding, and one factor was a meaning factor, including text comprehension as well as vocabulary knowledge. Perfetti and Hart also found that less-skilled readers had an additional (third) important factor; for those readers, two separate factors emerged for form knowledge: one for orthographic knowledge, and one for phonological knowledge. The authors concluded that for less-skilled readers, the link between representations of orthography and phonology is weak, and that overall, less-skilled readers can be thought to have more dissociated representations of components of word knowledge. More-skilled readers, on the other hand, have highly interconnected knowledge of the components of words, resulting in unified word representations.

Verhoeven and Van Leeuwe (2008) confirmed the relationship between word level skills and comprehension longitudinally with Dutch students, finding that vocabulary knowledge and efficient decoding of word forms predicts the development of reading comprehension. A more causal relationship between vocabulary instruction and comprehension and speed gains were found through controlled experiments by Beck, McKeown and colleagues (Beck, Perfetti, & McKeown, 1982; McKeown, Beck, Omanson, & Perfetti, 1983) (and see Stahl & Fairbanks, 1986, for a review). Furthermore, vocabulary knowledge predicted comprehension ability in young adult readers (aged 16-24) above and beyond decoding skill and general language comprehension (Braze, Tabor, Shankweiler, & Mencl, 2007).

Based on these analyses, two relevant concepts emerge. (1) Comprehension skill is highly related to vocabulary knowledge and the quality of lexical representations, but (2) even for skilled readers, and especially for less-skilled readers, there are dissociable word level skills. Form knowledge and vocabulary meaning knowledge account for different components of variability, with vocabulary meaning knowledge being more highly related to comprehension ability. However, as stated in Perfetti's (2007a) overview of lexical quality, "It remains an important goal of the Lexical Quality Hypothesis to show more specific consequences of the various components of lexical knowledge."

1.2.2 Lexical Quality and Fluent Reading

Speed is a component of both high quality lexical representations and skilled reading more generally. The ability to correctly identify words rapidly is considered an indicator of, if not part of the definition of, lexical quality. Although not all studies supporting the link between vocabulary knowledge and comprehension ability use timed tests or consider the speed component, some do. Perfetti and Hart (2002), for example, found that skilled comprehenders were faster to access word meanings, as evidenced by earlier signs of homophone confusion and subsequent meaning resolution. They use this finding as evidence that skilled comprehenders have higher quality lexical representations than less-skilled comprehenders. Likewise, in the factor analyses discussed in the previous section, both Hart (2005) and Landi (2005) used timed tests for the vocabulary and comprehension evaluations. Thus, their findings are actually that fast, accurate vocabulary access is related to fast, accurate reading comprehension.

Thus, it is important to consider not only accuracy in comprehension, but fluency in the *path* to comprehension. Two readers may achieve the same level of comprehension, but one reader may need to pour over the text, re-reading it several times, whereas the other reader may effortlessly understand the text on the first pass. In fact, even if both readers score the same on a set of comprehension questions following the text, the faster reader may have understood the text more deeply, overlooked by shortcomings in the evaluation method. More fluent readers should achieve better comprehension, because more of their cognitive capacity is freed from low-level word processing and is available for higher level consideration of the text (LaBerge & Samuels, 1974; Nathan & Stanovich, 1991), and correlational data suggest that this is the case (Pinnell et al., 1995).

This "fluency" skill – the ability to quickly and automatically identify words during reading – is often examined orally in younger children as a measure of reading ability or literacy gains (for a review, see Fuchs, Fuchs, Hosp, & Jenkins, 2001). Less fluent readers need to switch their limited attention between word identification and comprehension processes because the dual task of attending to both simultaneously exceeds the readers' attentional capabilities. This results in labored reading and weak comprehension. Fluent readers, on the other hand, are able to identify words almost effortlessly, and can keep a steady focus of attention on comprehension of the text. Although studies addressing fluency evaluation and instruction have generally had an emphasis on learning readers, fluency remains a differentiator of reading ability as late as high school (Rasinski et al., 2005).

1.2.3 Experience Improves Lexical Quality and Fluency

Educational studies have shown that the most effective way to achieve gains in fluency and vocabulary knowledge is through practice reading (Samuels, 2002). Cunningham and Stanovich (1991) found a positive relationship between print exposure and vocabulary in children even after general ability and decoding ability were accounted for, and it remained an important predictor of vocabulary in college students (Stanovich & Cunningham, 1992;

West & Stanovich, 1991). In fact, print exposure was a better predictor of vocabulary in older adults and college students than age was, even after controlling for working memory, general ability, and education level (Stanovich, West, & Harrison, 1995). Not surprisingly, instructional studies show similar results: more exposures to words results in better learning of the words (e.g Jenkins, Stein, & Wysocki, 1984; McKeown, Beck, Omanson, & Pople, 1985).

One mechanism for improvement in speed through practice is the unitization of practiced words. As children gain experience reading, there is a shift from letter-by-letter reading to gradually more unitized word representations, as illustrated by the gradual decrease in the importance of word length in determining semantic categorization reaction times from 2nd to 6th grade (Samuels, LaBerge, & Bremer, 1978). Unitization is akin to the idea of "chunking": text processing units become larger as words become more familiar. This shift is specific to the words that have been practiced, so that even adults revert to smaller visual processing units for unfamiliar words. For example, when adult readers were asked to read visually unfamiliar mirror-image words aloud, the effect of word length decreased with practice, indicating a similar shift to more unitized visual word representations of the practiced items (Samuels, Miller, & Eisenberg, 1979). Readers also find it more difficult to detect letters within more familiar words (frequently occurring, correctly spelled, grade-level appropriate, etc.) providing further evidence that these words are processed in larger units than less familiar (infrequent, misspelled, etc.) words (Greenberg, Healy, Koriat, & Kreiner, 2004; Healy, 1976, 1980, 1981; T. Cunningham, Healy, Kanengiser, Chizzick, & Willitts, 1988).

In the framework of the Lexical Quality Hypothesis (Perfetti & Hart, 2002, 2001), more exposures to words increases the robustness and specificity of the lexical representations, strengthens the connections between the learned words and the rest of the reader's knowledge and lexicon, and strengthens the connections between form and meaning and between orthography and phonology. All of these factors contribute to making the process of word identification faster and more accurate.

1.2.3.1 Neural Basis of Learning from Experience Functional MRI studies have contributed to our understanding of word learning by identifying neural changes resulting from repeated exposures to words. These changes correspond behaviorally to vocabulary learning and increased efficiency in word identification. Often, improved neural computation is manifested in fMRI as suppression effects of learning – an overall reduced amount of neural activity with expertise. Suppression of neuronal firing may be due to (a) a kind of "gain control" designed to help new information stand out, mostly in the cases of short-term repetition, (b) a sharpening of neuronal "tuning", whereby irrelevant or unnecessary neuronal firing is reduced, (c) a shortening of the time course of neuronal activity, in which communication simply occurs faster and firing stops sooner as the stimulus identity is settled upon more quickly, or (d) some combination of these factors (Grill-Spector, Henson, & Martin, 2006).

Katz et al. (2005) found this type of repetition reduction in an fMRI study in which new or repeated words were presented to participants in a lexical decision and naming task. As behavioral performance improved with familiarity, brain activity in regions associated with lexical processing (inferior frontal gyrus, supramarginal gyrus, and a posterior occipitotemporal region) decreased. Repeated words in the study were presented only three times for 1 second each time, so whether the effects are similar in nature to effects arising from a lifetime of experience with a word is questionable. However, Sandak et al. (Sandak, Mencl, Frost, & Pugh, 2004) found the same type of activity reduction in the same reading areas following a more extensive phonological non-word familiarization task and subsequent fMRI task, supporting the idea that such a reduction of activity as a reflection of processing efficiency is a defensible explanation for word familiarity, if not more robust learning.

Increased efficiency in the form of reduced neural activity is not the only way in which experience can shape neural circuitry. Changes in circuitry used for the task or increased connections between areas are other ways Posner et al. (1997) hypothesized skill could develop. In their 2004 study, Sandak et al. found changes of this sort following a semantic training paradigm in which participants learned meanings of non-words and then performed a simple naming task in the scanner. Unlike the phonological training which reduced neural activity in decoding areas, the semantic training increased neural activity in bilateral superior and middle temporal gyri. Such an increase in activity might be a result of increased connections to semantic areas with the semantic training. Although both the phonological training and semantic training used in the study resulted in faster and more accurate performance on relevant tasks, the underlying changes in neural circuitry were different depending on the type of training and the sub-process of reading that was becoming more accurate and efficient. Semantic familiarity seemed to result from increasing connections and activity, whereas phonological familiarity resulted from decreased activity (and perhaps increased efficiency) in decoding regions.

In summary, the rapid and automatic identification of words depends on the individual reader's prior experiences with the particular word being read, and how these experiences shape neural processing, regardless of whether the reader is learning to read or an adult. Word frequency effects, in which more frequently encountered words are understood more efficiently, can be thought of as an extension of this idea, with word-by-word "fluency" improving over our lifetimes of word experience.

1.2.4 Experience as an Individual Difference

Reading experience, and the number words an individual encounters, is a very large source of variability among individual readers. In children, at least, this may be due largely to differences in reading skill and the corresponding differences in reading speed. Biemiller (Biemiller, 1977) reports rapidly increasing differences between his most and least able reading groups in terms of the number of words read in a session of classroom reading, primarily because the least skilled readers were slow readers. In addition, Allington (1984) observed that less-skilled readers in the classroom were instructed in decoding at the expense of silent reading, resulting in less classroom exposure to word reading. This feeds the negative side of the reciprocal relationship between reading experience and vocabulary knowledge whereby readers with poor vocabulary knowledge are slower readers and therefore do not gain as much experience, making it difficult to improve vocabulary knowledge (see Stanovich, 1986). In contrast, readers with early success in reading are more likely to sustain a habit of reading through at least high school (A. Cunningham & Stanovich, 1997). Much of the work of Stanovich and colleagues has centered on the variability between individuals of all ages in the amount of exposure they have to print, and the large consequences of these differences on reading ability and vocabulary knowledge (see Stanovich & Cunningham, 1992).

1.2.5 Ability to Learn from Experiences as an Individual Difference

One important point to consider is that the quality of lexical representations does not depend solely on the frequency or even quality of experiences with that word. Studies have shown that different readers benefit to different degrees from the same experiences with words. More-skilled readers learned more after instruction followed by reading in context than less-skilled readers did from the same instruction and contexts (Jenkins et al., 1984), and among average and above average readers, a similar trend was seen that did not reach significance (Nagy, Herman, & Anderson, 1985).

Perfetti, Wlotko, and Hart (2005) taught more-skilled and less-skilled undergraduate comprehenders the meanings of very rare words for 50 minutes. Behaviorally, they found that although both groups started with equal (and low) knowledge of the very rare words, the more-skilled comprehenders were more accurate in a post-learning meaning judgement task. The implication is that the more-skilled comprehenders were better able to take advantage of their experience with the rare words during learning, resulting in a higher level of lexical quality from the same experience. This behavioral observation was confirmed in the ERP record. More-skilled comprehenders showed a more robust P600 episodic memory effect than did the less-skilled comprehenders. In addition, less-skilled comprehenders were not as facilitated in reading a semantic associate after viewing the semantically-related learned word. This was evidenced by a smaller reduction in the N400 component, related to meaning expectation or ease-of-integrating, for the associate. Together, the ERP results imply that the learning experience had less impact for the less-skilled comprehenders.

Nelson, Balass, and Perfetti (2005) found congruent results. Individuals were trained on the meanings of very rare words either visually or auditorily, but instead of having an equal time for learning, participants learned the words to criterion. This meant that the slower learners had more exposure to the words than the faster learners. Despite fewer exposures, the faster learners were still more accurate and more confident in later identifying the learned words. Specifically, these learners were less dependent on modality congruency between learning and test. If they learned the words visually, the faster learners were still able to recognize the word if it was presented auditorily, and vice versa. The slower learners suffered more from a change in modality at test than did the faster learners. This study further supports the idea that readers are differentially able to take advantage of exposures to words. In addition, it supports a key prediction of the Lexical Quality Hypothesis: skilled readers have more integrated form representations for words. Based on this study, this observation appears to be due to the reader's knowledge of grapheme-phoneme correspondences rather than based solely on exposure to the words in multiple modalities; in training, participants were only exposed to one modality, but they were still able to recognize the word when it was later presented in the alternative modality.

The conclusions were further supported in an ERP experiment (Balass, Nelson, & Perfetti, 2010) that showed a recognition memory effect (the P600) for reading trained words only in skilled readers. There was a particular skill difference in the phonologically trained condition (in this case the only condition tested in an incongruent modality).

Besides the comprehension skill differences often used to distinguish more-skilled from less-skilled adult readers, other reader skill differences may influence how easily the reader can learn from experiences with words. By the self-teaching hypothesis (Share, 1995), word learning from independent reading requires readers to be able to successfully re-code the written word into it's phonological form, thus allowing mapping from the visual form to the lexicon. On an item-by-item basis, words eventually become "lexicalized," at which point decoding is less important. Learners who are skilled decoders and have better phonological memory will be better equipped to establish strong orthographic word representations from fewer exposures to the word. A secondary component to the self-teaching hypothesis is that visual-orthographic memory will also distinguish readers in their ability to form new word representations.

These studies establish that important sources of individual variability in lexical representations include not only how often words have been experienced, but also the skills that allow a reader to benefit from those experiences. Evidence suggests that important skills include decoding, orthographic memory, and comprehension ability.

Overall, studies examining the relationship between reader skill and the quality of lexical representations show that (1) comprehension skill and fluent reading depends on lexical quality, and (2) the quality of word representations comes from a combination of the amount of experience a reader has with a word and the relevant skills that enable that reader to take advantage of experiences with words.

1.2.6 Eye movements as a Measure of Reading Skill

Although much of the background presented so far has focused on efficient, accurate lexical access as measured through vocabulary and fluency assessments, many studies of adult reading have obtained fine-grain word-by-word information about the speed and ease of text processing through tracking readers' eye movements while they read silently. There are two major benefits to using eye-tracking to understand the cognitive processes underlying reading: (1) eye-tracking provides precise temporal-spatial information, recording exactly where the eyes are fixated in a text and for how long, and (2) eye-tracking studies allow linear displays of text, simulating a normal reading situation in which participants can read backward and forward in the text with no additional task response required. However, the validity of eye movement studies depends on there being an "eye-mind link," a relationship between cognitive processing and eye movements (see Schooler, Reichle, & Halpern, 2004, for a discussion of this idea). Most researchers investigating eye movements during reading agree that there are both occulo-motor factors and cognitive factors that contribute to decisions about when and where to move the eyes. Much current debate revolves around the exact nature of contributions from these factors, resulting in several models generated to test the quantitative predictions of various theories (Reichle, Pollatsek, & Rayner, 2006, and see the entire January 2006 special issue of *Cognitive Systems Research*). However, most eye movement research being carried out assumes that there is a tight relationship between the location of fixations and the content being processed, and between the duration and number of fixations, and how difficult processing is. In this way, we can learn which text factors, lexical factors, and individual reader differences influence reading behavior. We can learn more specifically when and in what way these factors influence reading by examining a variety of distinct measures gleaned from the eye movement record. These include the durations of initial fixations on words, various cumulative measures of viewing time when multiple fixations are made on words, whether a word was skipped, whether it was re-fixated on the first pass, and whether a regression was made back to the word. The general pattern of data across many studies is that more difficult text as well as less-skilled reading results in patterns of longer, more frequent fixations, more regressions, and slower reading times.

1.2.7 Influence of Text Factors and Lexical Factors on Efficient Reading

Reading difficult text, whether at the level of the single word, sentence, or entire text, results in elements of the typical "difficult reading pattern": longer fixations, more fixations, shorter saccades, more regressions, and slower reading times. Rayner (1986) also found that the perceptual span, the size of the window of text from which useful information is being gleaned during any fixation, shrinks when text is more difficult. This implies that attention is more focused on the foveated region when there are high processing demands, and provides evidence of the hypothesized relationship between the size of a processing unit and the difficulty of the text, with more difficult text resulting in smaller processing units.

Various properties of the individual words making up the text have been found to affect processing demands as measured by eye movements, replicating effects found in behavioral paradigms. These include word length, predictability, frequency, lexical ambiguity, age of acquisition, subjective familiarity, concreteness, phonological neighborhood size, orthographic neighborhood size and orthographic neighborhood frequency, among other properties (for example, Andrews, 1997; Juhasz & Rayner, 2003; Pollatsek, Perea, & Binder, 1999; Rayner, Sereno, & Raney, 1996; C. C. Williams, Perea, Pollatsek, & Rayner, 2006; Yates, Friend, & Ploetz, 2008). Ultimately, these kinds of lexical factors do not necessarily reflect properties that are intrinsic to the word. For example, Staub and Rayner (2007) differentiate between "intrinsic" lexical factors (e.g. frequency, morphology, and lexical ambiguity) and "relational" lexical factors, which have to do with how a word fits in its context (e.g. predictability, semantic priming, plausibility in context). But even Staub and Rayner's category of "intrinsic" properties can be broken down further. Some are independent of the context, the reader, and the reader's lexicon (such as word length), some are dependent on reader experience (such as frequency or familiarity), and some are dependent on the reader's lexicon (such as orthographic neighborhood size, defined as the number of words that are only one letter different from the target word). Although the concept of neighborhood size, for example, exists without reference to a particular reader, if a particular reader doesn't actually know a word's neighbors, those neighbors should not affect that reader's lexical processing. In addition, the degree of orthographic specificity a reader has for words can affect the relationship between words and their neighbors, with better spellers or more advanced readers showing neutral or inhibitory priming effects from neighbors, rather than the facilitatory effects seen in weaker spellers or developing readers (Castles, Davis, Cavalot, & Forster, 2007; Andrews & Hersch, 2010). Likewise, although there are external word frequency measures, they are estimates of how frequently readers are actually likely to encounter words, and without that relationship to the reader's experience, the variable is meaningless for understanding lexical processing. This leaves few factors that are actually "intrinsic", with some factors being perhaps better categorized as "knowledge" factors (dependent on an individual's knowledge, lexicon, or experience), and some being, as Staub and Rayner (2007) proposed, "relational" factors that depend on the context in which they appear.

Although relational factors are generally studied by controlling the kinds of context in which words appear, "knowledge" factors are generally studied by using a proxy for the average reader's experience or knowledge based on corpus measures such as word frequency or number of neighbors. This has been an effective strategy for understanding the average reader, often made even more effective when subjective measures, such as familiarity ratings, are employed (as will be explored in the next section). However, as discussed in previous sections, individual readers' actual experiences, their ability to learn from their experiences, and their resulting word knowledge are quite variable, and these differences are a good starting point in understanding individual differences in reading skill among adult readers.

1.2.7.1 Frequency and Familiarity As An Example The case of word frequency effects provides a good example of when it is useful to think of lexical properties as "knowledge" factors, even when trying to understand the average reader. It is fairly non-controversial that the frequency of our encounters with words influences the ease with which we process those words. Frequency effects have been widely reported in the word recognition literature in a variety of tasks including word reading and visual lexical decision (Balota & Chumbley, 1984; Forster & Chambers, 1973; Frederiksen & Kroll, 1976; Monsell, Doyle, & Haggard, 1989; Schilling, Rayner, & Chumbley, 1998), picture naming (Oldfield & Wingfield, 1965), tachistoscopic recognition thresholds (Broadbent, 1967; Broadbent & Broadbent, 1975; Solomon & Postman, 1952; Winnick & Kressel, 1965), eye movements (Schilling et al., 1998), and spoken word recognition (Connine, Mullennix, Shernoff, & Yelen, 1990; Dahan, Magnuson, & Tanenhaus, 2001; Marslen-Wilson, 1987).

However, Gernsbacher (1984) demonstrated that subjective familiarity ratings are a better single predictor of reading performance than objective frequency counts. Previous literature had shown mixed results for interactions between frequency and bigram frequency, concreteness, and polysemy, especially in the low-frequency range. For example, in the low frequency range, Rice and Robinson (1975) found that participants were faster to respond to low bigram frequency words than high frequency bigram words, whereas Biederman (1966) found exactly the opposite. Gernsbacher's hypothesis was that familiarity was really the main factor, and that although frequency approximated familiarity in the high frequency range, words in the low frequency range randomly varied in familiarity between conditions and between studies. She demonstrated that for the bigram frequency studies, stimuli used in the previous experiments varied in familiarity in such a way as to explain the pattern of inconsistent results. For example, in the low-frequency range, more of Rice and Robinson's low bigram frequency words were highly familiar than their high bigram frequency words, whereas the opposite was true in Biederman's study. When Gernsbacher repeated the lexical decision studies but varied familiarity instead of frequency, she found no effects of bigram frequency, concreteness, or polysemy. In addition, the familiarity ratings correlated strongly with a measure of performance combining latency and accuracy, (rs = -.86, -.89, -.87, and -.78 for the four studies) and accounted for 71% of the variance in performance across all words and participants in the studies. Previous studies could not account for that much variance when objective frequency measures were used. Howes and Solomon (1951) found that frequency could account for an average of 50% of the variance in tachistoscopic thresholds, and Whaley (1978) could similarly account for only 46% of the variance in a lexical decision task. Gernsbacher's explanation of familiarity's strong ability to predict performance was that subjective familiarity ratings were a more reliable measure of true experiential frequency, and that they might also take into account the number of times a word is spoken, written, or heard in addition to how often it is read. She proposed that familiarity ratings may be performed by tapping into a stored representation of frequency information for each word, a memorial frequency record that is independent of a words other attributes.

More recently, Kacinik, Shears, and Chiarello, (2000) found that familiarity was more highly correlated with performance in a lexical decision task (r = -.70) than was objective frequency (r = .22). Familiarity was correlated with imageability, noun-verb distributional distance (NVDD, a measure of the typicality of the context in which a noun or verb appears compared to other nouns and verbs), Francis-Kucera frequency (1982) and Usenet frequency (extracted from newsgroup conversations on Usenet), r = .22, .23., .39, and .40 respectively. However, familiarity was still able to account for unique variance in reaction time for the lexical decision tasks. The authors urge experimenters to consider the importance of processing differences based on the familiarity of stimuli to the subject population instead of controlling primarily for frequency and imageability.

If we acknowledge that word frequency is an estimate of how often readers have encountered a given word relative to other words, and that the reason more-frequent words are read more efficiently is that experience with words breeds high quality lexical representations, then using a subjective measure may have more predictive power because it is more likely to directly reflect the quality of readers' lexical representations. Whether this is because familiarity ratings are a more accurate measure of true multi-modal experiential frequency than are objective frequency counts, as Gernsbacher (1984) suggests, or because readers are actually basing their familiarity ratings on how easily they feel they can identify the word, is an unanswered question.

It is possible that the familiarity ratings are not tapping into a stored bit of lexical frequency information, but instead are based on the pattern of neural activity associated with processing the word. If the neural activity shows the signature pattern of non-novel processing (e.g. processing is fast, highly tuned, or well-connected), a participant might use this information to make familiarity ratings. Instead of the ratings reflecting an independent lexical property (frequency) that influences neural processes of word recognition, the ratings may be measuring something about the neural processing itself. In that case, familiarity ratings would depend on the same factors (such as frequency) that cause learning, and would at least partially measure lexical quality rather than simply estimate frequency of exposure. This would be true especially when readers are asked to rate "familiarity" rather than frequency.

1.2.8 Summary of Background

Previous research has identified that vocabulary knowledge, including the speed and ease with which lexical representations can be accessed, is an important component of skilled reading and comprehension. A reader's ability to fluently and accurately read a word depends on the quality of the lexical representation, which in turn depends on both how much experience the reader has had with a word, and how well the reader is able to establish and refine their lexical representations with each exposure. Thus, any observed reading behavior is a result of the combination of properties of words in the text being read and the reader's current knowledge of the words in the text.

1.3 OVERVIEW OF STUDIES AND RESULTS

The purpose of this work is to understand how individual differences between adult readers, as well as differences in an individual's knowledge of specific words, manifest in moment-to-moment reading processes through the examination of eye movements during reading. This was achieved through a factor analysis and three experiments.

1.3.1 Study 1: Individual Differences Between Adult Readers: Factor Analysis of Adult Database

In Study 1, a factor analysis was performed to identify orthogonal dimensions of variability among adult readers. Rather than choosing a single dimension or individual (perhaps correlated) tests to define reading skill, we took a data-driven approach to defining functionally distinct dimensions of variability. Participants took a battery of tests relating to various aspects of intelligence, reading skill, and reading history, and a factor analysis was performed to create new axes of variability. Five main dimensions emerged from the analysis reflecting: (1) speed and reading experience (expertise), (2) sublexical skills, (3) accuracy, (4) learning/memory, and (5) amount of "casual reading" (magazines, newspapers, and internet).

1.3.2 Study 2: Relationship Between Individual Differences and Eye Movements

The goal of Study 2 was to characterize how the dimensions of individual variability defined in the factor analysis contributed to differences in word reading behavior. This was achieved by tracking readers' eye movements while they read paragraphs of text for meaning. Analysis was done on an individual word and an individual reader basis, which allowed us to understand not only the overall effects of reading ability, but also the interaction between reading ability and lexical characteristics (both "intrinsic" and "knowledge" characteristics as discussed previously).

Results showed large benefits for readers scoring highly on the expertise dimension, especially for less-frequent words, supporting the notion that readers with more exposure to text have had more opportunities to experience less-frequent words and subsequently build high quality lexical representations for these words. Sublexical skills of the variety that might be important for self-teaching expedited the early fixations on words (those probably more reflective of word form processing), especially for more frequent words. This finding supports the hypothesis that in adults, on-line decoding of less frequent words is less of a differentiator, but that good sublexical skills enable readers to achieve greater success in lexicalizing or unitizing frequent words, resulting in faster identification of frequent word forms.

1.3.3 Study 3: Effects of Partial Word Knowledge on Eye Movements (Training Paradigm)

The goal of Study 3 was to to characterize the relationship between word knowledge (orthographic, phonological, and meaning components) and reading behavior more directly, avoiding proxies for measures of word knowledge such as frequency or familiarity ratings. This was achieved by employing a rare word training paradigm to control which components of knowledge a reader learned about a word (orthographic, phonological, and/or meaning), and the quality of that knowledge (by varying number of exposures). However, given that different readers vary in how well they can learn from exposures to words, we could not assume that readers acquired similar degrees of lexical quality after the training session. Because of this, we also accounted for the individual skill differences identified in the Study 1 factor analysis.

Results showed that for these previously unknown words, there was no main effect of reader expertise, but more expert readers benefitted more from the training. Good sublexical skills, on the other hand, supported reading for the newly trained words regardless of the training condition, in addition to providing better learning from the addition of auditory information specifically. This suggests that sublexical skills are generalizable to new words and are utilized when reading new words, whereas reading experience helps with the specific information that has been previously experienced. Better learning and memory skills resulted in better reading performance, an expected result given that this was a training paradigm.

Additionally, results showed that the individual components of word knowledge affected the pattern of reading behavior in different ways. Early measures were primarily affected by visual and auditory (form) training, corresponding to orthographic and phonological knowledge, whereas later measures were more affected by auditory and meaning training, corresponding to phonological and semantic knowledge. These findings suggest that lexical knowledge is multidimensional, with different components of knowledge established somewhat independently via different types of exposures and via the skills readers have to capitalize on each encounter. More nuanced results are discussed in Chapter 4.

1.3.4 Study 4: Effects of Partial Word Knowledge on Eye Movements (Testing Paradigm)

Study 4 takes a different approach to achieving the same goal of characterizing the direct relationship between components of word knowledge and reading behavior. Rather than trying to *control* the type and quality of word knowledge readers have, as in Experiment 2, we instead attempted to *measure* the type and quality of word knowledge readers have. Participants who read paragraphs in Study 2 were brought back to test their spelling, pronunciation, and meaning knowledge of certain target words within the paragraphs. It was assumed that their knowledge of the words remained about the same between the two sessions, and we could therefore relate their scores on the vocabulary test to their previous reading behavior. This approach removed the complication of participants having just learned the words and perhaps trying to recall their training during reading.

Results replicated many of the findings of Study 3. Phonology effects inflated first-pass refixation durations, and meaning effects were seen in later measures including re-reading durations and total viewing times. Orthographic uncertainty showed an interesting pattern in which less orthographically well-known words had inflated probabilities of re-reading and longer total viewing times, and the least certain words also had *faster* first pass reading times. It may be beneficial for readers to skim unknown words on the first pass through the text to maintain memory for the meaning being constructed from the sentence or passage. However, it is beneficial to eventually look back at the word to determine if the meaning is known, and if not, to lay the groundwork for the new lexical representation with whatever information is available.

2.0 STUDY 1: INDIVIDUAL DIFFERENCES BETWEEN ADULT READERS: FACTOR ANALYSIS OF ADULT DATABASE

2.1 INTRODUCTION

The goal of Study 1 was to identify and characterize dimensions of variability reflecting individual differences between adult readers. We wanted to depart from an a priori definition of adult reader skill, often defined by scores on a comprehension test, and instead use a data-driven approach to identify multiple continuums of skills, knowledge, or experience along which adult readers vary. To understand the structure of variability in reading ability, we used a large database of 2,123 readers who took a battery of tests designed to measure skills in a variety of reading-relavent areas including decoding, spelling, phonological awareness, comprehension, vocabulary, print exposure, and for a subset of the readers, reading history. We computed the factor structure of this database to understand which tests shared variance and may be drawing on the same underlying skills (or may even be more causally linked). The resulting factors were then used as the dimensions of reading skill defining individual differences in Studies 2 and 3.

We expected to replicate results of similar PCA's performed by Hart (2005) and Landi (2005) by finding a dissociation between form-related spelling/decoding skills and meaning-related/comprehension skills or knowledge. However, this study also extends prior analyses through changes in the scoring methods of timed tests and the addition of several assessments. Instead of a single composite score, timed test scores were broken into a speed component (% of questions attempted) and an accuracy component (% of attempted questions that were answered correctly). This allows a separation between readers who score poorly in a composite test score because they are both slow and inaccurate, those who score poorly because they are slow, despite being accurate, and those who score poorly because they are inaccurate, despite being fast. This study also includes scores from assessments not included in previous studies, including scores from a phonological awareness test, a test of non-verbal intelligence, and a reading history questionnaire (a complete description of all of the tests follows).

2.2 METHODS

2.2.1 Participants

Two thousand one hundred twenty-three participants from the University of Pittsburgh community took a battery of tests designed to assess reading skills (described in more detail below). Most participants were undergraduates from the University of Pittsburgh, although the sample included some participants from the greater community. The mean age of participants was 18.87 years. 37.8% were males and 59% were females (the remaining 3.1% were missing this information). Participants either received introductory psychology course credit or payment for their participation. All participants completed the tests between the fall semester of the 2005-2006 academic year and the summer (2009) semester of the 2008-2009 academic year. A subset of 1450 participants who took the tests between spring of 2007 and summer of 2009 also took the Adult Reading History Questionnaire, and a further subset of 90 participants between the summer of 2008 and the spring of 2009 also received a battery of working memory tests, described below.

2.2.2 Procedure

All participants gave informed consent prior to taking the following pencil-and-paper tests:

Phonological Awareness Test (PhAT) The PhAT tests phonological awareness by requiring participants to manipulate phonemes within words. Participants start by removing phonemes from words; for example, they are asked to remove the /d/ sound from "middle," resulting in the word "mill." This requires participants to pay attention to the way the word sounds, not the way it is spelled (in this example "mile" is an incorrect answer). Then, the participants are asked to add a sound where the previous phoneme was removed. In this example, they now insert an /s/ sound where the /d/ had been removed to make the word"missile." Participants are scored based on whether the answer sounds correct ("mil" and "mill" would be correct in the first example), and whether the answer was a correctly spelled real word, as participants are told it should be ("mill" or "mile" would get this point, but not "mil"). If both points are obtained, it means the answer was correct, otherwise partial credit is given. See Appendix C for the full test.

Ravens Progressive Matrices The Ravens Progressive Matrices is a measure of non-verbal intelligence. Each test item is a three-by-three array of patterns with the final patten of the nine omitted. Participants must choose from among six choices the pattern which completes the series. This test was administered as a 15 minute timed test, and participants were instructed not to skip any items. Each item is progressively more difficult than the last (Raven, 1960). This test is scored separately for speed (% of items answered) and accuracy (% of answered items that were correct).

Author Recognition Test (ART) The ART (Stanovich & West, 1989) functions as an estimate of print exposure. It consists of a checklist of 80 names of real people, some of whom are authors. Participants are instructed to place a check mark next to any names they recognize as being popular writers of books, magazines, or newspaper columns. The test is scored using d'. The validity of the ART as an indicator of print exposure has been documented in several studies illustrating significant positive correlations with orthographic processing (Stanovich & West, 1989), and comprehension ability (West, Stanovich, & Mitchell, 1993). The version that we presented is modified such that some of the foils (e.g. published psychologists) were replaced with foils who were more clearly non-authors. See Appendix B for the full test.

Nelson-Denny Comprehension Test The Nelson-Denny Comprehension Test (N. J. Nelson & Denny, 1973), Form E (Brown, Bennett, & Hanna, 1981) consists of eight passages, each followed by 5-answer multiple choice comprehension questions about the passage for a total of 36 items. Participants have 15 minutes to complete as much of the test as they can (instead of the usual 20 minutes). The test is scored for both speed (% of items completed) and accuracy (% of answered items that are correct).

Nelson-Denny Vocabulary Test The vocabulary portion of the Nelson-Denny reading test (Brown et al., 1981) is given as a 7.5 minute timed test (half of the normal time allotted), and participants are instructed not to skip any of the items, which get progressively more difficult. The test is a multiple-choice test in which participants choose each word's definition from 5 choices. Questions are presented in a complete-the-sentence style (e.g. A *brochure* is a...) There is both a speed (% of items completed) and accuracy (% of completed items correct) measure.

Real Word Test Participants read a list of non-words and select the ones that sound like real words (those that are psuedohomophones). Items on the test were adapted from items used by Olson and his colleagues on a test of orthographic knowledge (Olson, Wise, Conners, Rack, & Fulker, 1989). More difficult items were added by Perfetti and Hart. The test is scored using d'. For the full test, see Appendix D.

Spelling Test Participants are shown a series of letter strings and are asked to identify items that are correctly spelled real words. A subset of items come from a test by Olson and colleagues (Olson et al., 1989). More difficult items were added by Perfetti and Hart, some of which were obtained from the Baroff Spelling Test(Perfetti & Hart, 2002). The test is scored using d'. See Appendix E.

Adult Reading History Questionnaire (ARHQ) The ARHQ is a series of questions surveying the participants' reading habits as well as any history of difficulties with reading, learning, spelling, and memory. The test is based on Lefly and Pennington's (2000) modification of Finucci's (1982; 1984) questionnaire. A question about internet reading was added to our version. Each question is answered using a scale from 0-4 (with .5 answers allowed). See Appendix A for the complete questionnaire.

2.3 RESULTS

2.3.1 Full Set of Participants

Means and standard deviations were computed for each of the assessment scores that were completed by the full set of participants (all tests except the ARHQ). See Table 1.

	Mean	SD
ART d’	1.54	.57
Real Word d’	2.16	.71
Spelling d’	1.95	.56
Vocabulary % Attempted	0.59	.14
Vocabulary % Accuracy	0.84	.11
Comprehension % Attempted	0.71	.16
Comprehension % Accuracy	0.82	.11
Raven’s % Attempted	0.92	.12
Raven’s % Accuracy	0.51	.22
Phonological Awareness	0.82	.13

Table 1: Descriptive Statistics for Reading Tests

Pearson’s correlation was also computed for the subset of tests that were completed by the full set of participants. See Table 2. Most of the measures were significantly but only moderately correlated with each other. Interestingly, although most measures were positively correlated, the speed (% attempted) on the Raven’s Matrices test was negatively correlated with all of the other test scores, except for the two other speed measures (for Nelson-Denny Comprehension and Vocabulary). It was most negatively correlated with the accuracy on the Raven’s, indicating that there was a speed-accuracy tradeoff for this test that was not present in the Nelson-Denny tests, in which speed and accuracy were positively correlated. Answering fewer questions on the Raven’s was, in fact, associated with higher accuracy scores on all tests.

Table 2: Correlations Between Reading Tests

	ART	Real Word	Spelling	Vocab %Attem	Vocab %Acc	Comp %Attem	Comp %Acc	Ravens %Attem	Ravens %Acc	PhAT
ART	--	.270**	.315**	.420**	.457**	.340**	.291**	-.018	.111**	.313**
Real Word	.270**	--	.285**	.125**	.307**	.036	.255**	-.117**	.206**	.468**
Spelling	.315**	.285**	--	.286**	.224**	.216**	.180**	-.007	.038	.307**
Vocabulary % Attempted	.420**	.125**	.286**	--	.209**	.637**	.176**	.205**	-.001	.209**
Vocabulary % Accuracy	.457**	.307**	.224**	.209**	--	.148**	.428**	-.147**	.238**	.348**
Comprehension % Attempted	.340**	.036	.216**	.637**	.148**	--	.080**	.259**	-.048*	.097**
Comprehension % Accuracy	.291**	.255**	.180**	.176**	.428**	.080**	--	-.110**	.243**	.278**
Ravens % Attempted	-.018	-.117**	-.007	.205**	-.147**	.259**	-.110**	--	-.316**	-.104**
Ravens % Accuracy	.111**	.206**	.038	-.001	.238**	-.048*	.243**	-.316**	--	.277**
PhAT	.313**	.468**	.307**	.209**	.348**	.097**	.278**	-.104**	.277**	--

* correlation is significant at the $p < 0.05$ level (two-tailed)

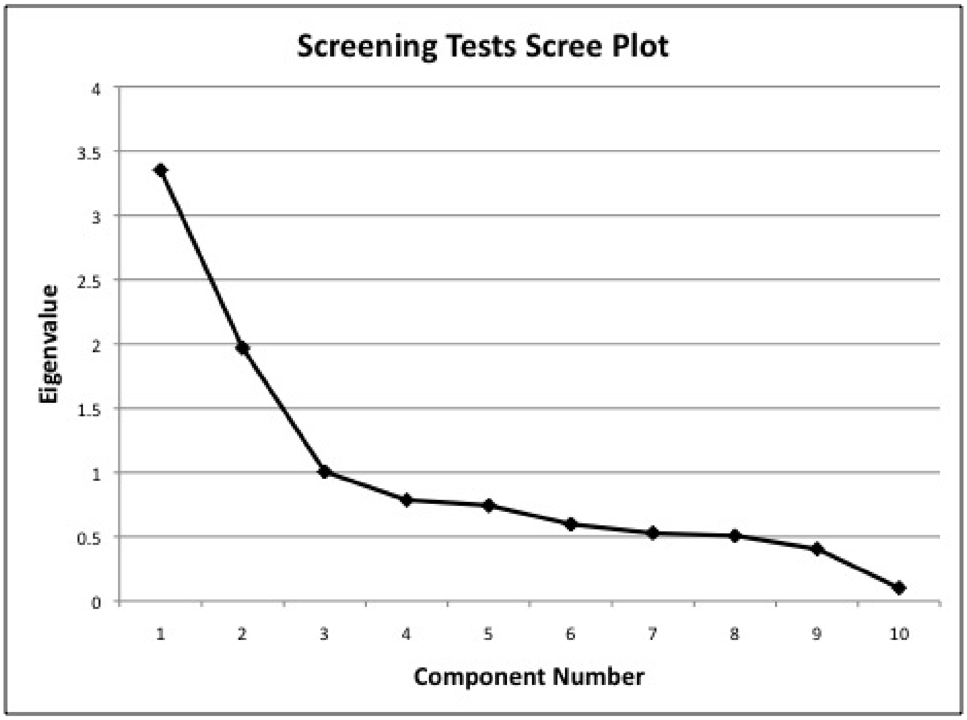
** correlation is significant at the $p < 0.01$ level (two-tailed)

Because we were interested in reducing the data to a few key dimensions along which adult readers vary, the correlation matrix was used in a principal components analysis (PCA). For all analyses, the resulting factors from the PCA were rotated to maximize the degree to which each original test score was correlated strongly with only one factor. This improves interpretability of the factors and enables the original test scores to be differentiated from each other or clustered with each other in terms of the set of factors. Because it is expected that various reading skills are correlated, the Promax rotation, which allows an oblique solution (Kaiser, 1958), was applied to the first two PCA's performed (the current PCA of the larger dataset and a PCA of the ARHQ test questions). A Varimax rotation, which provides an orthogonal solution, was applied to the final PCA to produce uncorrelated dimensions of individual differences for use as independent variables in studies 2 and 3. This eliminated problems of collinearity in the analyses for those studies.

First, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was computed for each variable to determine whether any measures should be excluded from the analysis. This statistic is an index between 0 and 1 where larger values indicate more shared variance with the rest of the variables, thus lending themselves well to factor analysis. All variables had an adequate KMO measure, with the lowest being .661 for the Nelson-Denny comprehension speed (% attempted). The overall KMO with all variables was .767, in the "middling" range in terms of factorial simplicity, according to Kaiser (1974) . Kaiser suggests that anything below .5 is unacceptable for factor analysis. Bartlett's Test of Sphericity was $p < .001$, indicating that a factor analysis was appropriate for this set of data.

We extracted the three factors that had Eigenvalues greater than one (See Figure 1). Collectively, these three factors accounted for 58.5% of the variance in the data. Factor 1 accounted for 30.3%, Factor 2 for 18.5%, and Factor 3 for 9.7%. The factor loadings (correlations with each factor) for these three factors were computed for each of the test scores. Scores were assigned to a factor if the loading was greater than .5 (see Landi, 2005). Factor loadings are shown in Table 3.

Figure 1: Scree Plot for Factor Analysis of Full Set of Participants



Loading heavily on the first factor were the Author Recognition Test as well as the speed measures for the Nelson-Denny vocabulary and comprehension tests. From this we can infer that Factor 1 measures a reading efficiency skill that is related to how much reading one does. One might think of this as reading expertise. Factor 2 comprises the accuracy measures for the Nelson-Denny vocabulary and comprehension tests, as well as the accuracy of the Raven's test of non-verbal intelligence. The speed of the Raven's loaded negatively on this factor, as might be expected from the correlation matrix and the fact that the test gets progressively more difficult. This factor, then, generally reflects accuracy on the higher-level tests. Participants with high scores along this factor proceeded through the timed tests at a pace that enabled them to maintain a high percentage of accurate answers. Thus, this component reflects a reader's emphasis on accuracy. The third factor seems to measure sublexical skill in reading, with the spelling test, phonological awareness test, and real word test (a decoding test) all loading heavily on it.

To recap, the factor analysis indicates that there are three main axes of variance in the data: one related to reading expertise, one related to accuracy emphasis, and one that

Table 3: Three factor PCA solution with Promax Rotation

	Factor 1 Expertise	Factor 2 Accuracy Focus	Factor 3 Sublexical
ART	0.630	0.449	0.476
Vocabulary % Attempted	0.847	0.099	0.289
Comprehension % Attempted	0 .831	-0.017	0.141
Vocabulary % Accuracy	0.373	0.692	0.450
Comprehension % Accuracy	0.283	0.669	0.336
Raven's % Attempted	0.380	-0.551	-0.087
Raven's % Accuracy	-0.081	0.690	0.204
Spelling	0.377	0.114	0.687
Phonological Awareness	0.212	0.451	0.769
Real Word	0.105	0.371	0.791

Shaded cells indicate a factor loading > .5

reflects sublexical, form level reading skill.

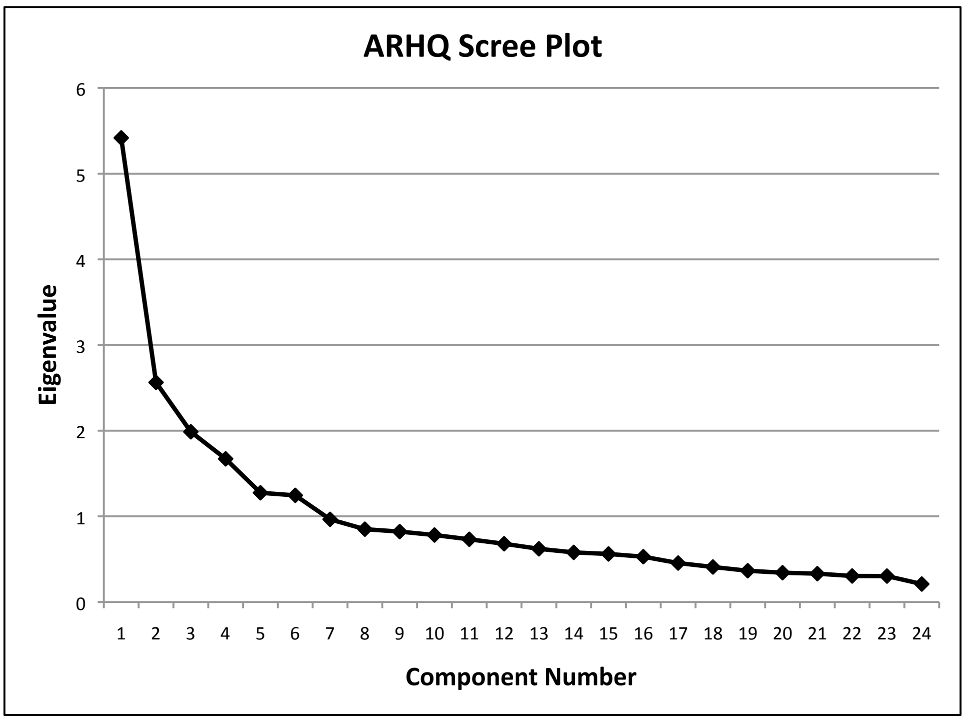
2.3.2 Participants who took the ARHQ

A subset of 1450 participants also took the Adult Reading History Questionnaire (ARHQ). We wanted to examine the factor structure of data for this group of participants to discover any additional sources of individual differences exposed by the questions on the ARHQ. The questions in the ARHQ address a broad spectrum of possible reading and learning difficulties as well as attitudes and habits toward reading and classwork more generally. Because there is such a wide range of topics covered by the test, an overall composite score would be difficult to understand as reflecting a single skill or attitude. Instead, we did a factor analysis of the ARHQ to group questions into sub-areas that would provide separate scores along the dimensions of reading and learning that the ARHQ covers. These sub-scores were computed so that they could then be entered into the larger factor analysis with the rest of the reading tests in place of a single overall ARHQ score.

For this dataset (the questions on the ARHQ), the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was .829, in the "meritorious" range for factor analysis (Kaiser, 1974), and Bartlett's Test of Sphericity was $p < .001$, indicating that a factor analysis was appropriate. A PCA with Promax rotation was performed, allowing correlations between the resulting factors.

Based on the elbow of the scree plot, as well as interpretability of the five vs. six factor result (See Figure 2), five factors were extracted. These cumulatively accounted for 53.8% of the variance in the ARHQ scores (22.6%, 10.7%, 8.3%, 7.0%, and 5.3%, respectively). Table 4 shows the factor loadings for each question on the test. For the full questions in non-abbreviated form, see Appendix A.

Figure 2: Scree Plot for Factor Analysis of ARHQ



The first factor included questions about attitude towards reading, books read for pleasure, and current reading speed. It is of note that similarly to the factor analysis of the battery of reading tests, reading speed seems to co-vary with amount of reading, even by self report. Factor 2 is related to spelling and learning skill. Questions loading on this factor included whether the participant had difficulty learning to spell and read, current spelling ability, reading skill as a child, whether extra help in reading was needed, and whether the participant struggled to complete their work in school. Factor 3 included questions that might indicate a reading disability. The questions probing whether there was difficulty learning to read also loaded highly on Factor 3, along with questions about whether the student had difficulty learning letter or color names, whether repeating a grade was considered, and whether the reader reverses letter or number order. Although the question asking about the reader's attitude towards school as a child did not load heavily on any factor, it was most closely correlated with this reading disability factor. Factor 4 contained only questions pertaining to what can be called "casual reading" – how often participants read daily papers, the Sunday paper, the internet, or magazines. It is interesting that these "casual reading" questions are a separate source of variability from book reading and general attitude about reading (Factor 1). It suggests that participants may have viewed the questions about attitude toward reading and reading for pleasure as meaning specifically book reading, and that there is more variability in how often people read books than how often they read casually (because Factor 1 accounts for the most variance). Alternatively, maybe only those with a positive attitude toward reading take the time to read books, whereas even those who consider themselves non-readers tend to read magazines and internet articles. Factor 5 is related to memory. Questions that loaded on the memory factor concerned difficulty remembering addresses, phone numbers, dates, names, places, and complex verbal instructions.

The factor analysis helped to break down the questions on the ARHQ into related groups such that the ARHQ can yield five sub-scores instead of one overall score. To recap, these sub-scores reflect the following five dimensions: reading attitude/book reading, spelling/learning, reading disability, casual reading, and memory. The five sub-scores were computed for each participant using the regression method available in SPSS. This method estimates normalized factor score coefficients for each participant.

Table 4: ARHQ five factor PCA solution with Promax Rotation

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
	Books/Attitude	Spell/Learning	Rdng Disab.	Casual Rdng	Memory
attitude towards reading	0.862	0.272	0.158	0.120	0.123
reading for pleasure	0.850	0.145	0.095	0.194	0.038
books for pleasure	0.830	-0.147	0.154	0.212	0.025
reading speed current	0.643	0.472	0.153	0.201	0.237
difficulty English classes	0.581	0.364	0.149	0.135	0.325
reading for work/class	0.381	0.198	0.243	0.145	0.151
difficulty learning to spell	0.096	0.795	0.329	0.105	0.260
current spelling	0.243	0.751	0.134	0.159	0.262
child reading skill	0.390	0.731	0.420	0.073	0.180
difficulty learning to read	0.252	0.703	0.587	-0.019	0.129
extra help learning to read	0.195	0.648	0.551	-0.045	0.074
struggle to complete work	0.279	0.482	0.450	-0.070	0.289
diff. learn letter/color names	0.185	0.325	0.743	0.017	0.146
consider repeating grades?	0.167	0.196	0.681	0.121	0.104
reverse letter/num order?	0.016	0.399	0.660	0.119	0.241
current rev. lett/num order?	0.056	0.330	0.532	0.186	0.312
school attitude as child	0.291	0.165	0.351	-0.023	0.205
daily newspapers	0.176	0.137	0.085	0.822	0.142
Sunday paper	0.223	0.130	0.055	0.814	0.145
internet reading	0.165	0.124	.104	0.645	0.045
magazines for pleasure	0.152	0.023	0.106	0.606	0.100
remember names/places	0.070	0.245	0.209	0.098	0.850
remember address/phone/date	0.064	0.203	0.186	0.086	0.847
remember verbal instructions	0.289	0.278	0.227	0.135	0.711

Darkly shaded cells indicate a factor loading > .5

Light shading denotes highest loadings for questions that do not load heavily on any factor

Before conducting the factor analysis with the ARHQ dimensions included, we first replicated the original factor analysis on this sub-group of participants to ensure that the the variability in this dataset followed a similar pattern to the variability in the larger group. Indeed, three factors came out of the analysis, and tests clustered together on factors in the same way as reported above in the larger group. The only difference was that the sub-lexical/form factor for this dataset accounted for more variance (18.2%) than the factor relating to accuracy/non-verbal intelligence (9.4%). In other words, factors 2 and 3 reversed order.

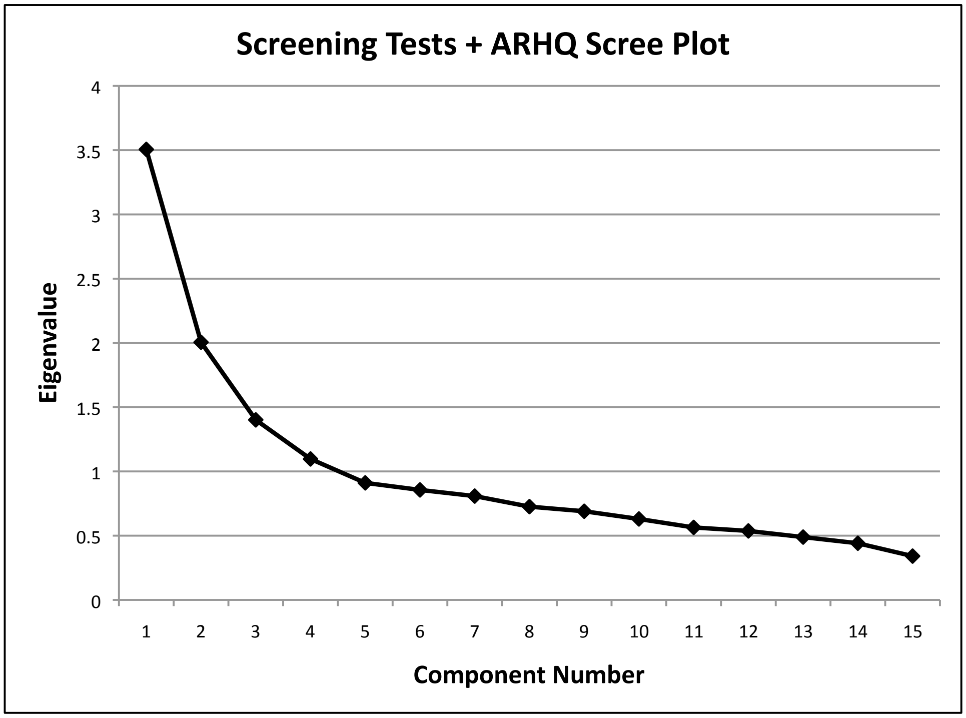
Because none of the clustering of test scores changed, this was considered sufficiently similar and a PCA with varimax rotation was performed on all of the original test scores plus the scores for the 5 dimensions of the ARHQ ($KMO = .790$; Bartlett's Test of Sphericity was $p < .001$). The varimax rotation was chosen in this case because these are the final factor scores that will be used as independent variables in Experiments 1 and 2. Having orthogonal factors eliminates the problem of collinearity between the independent variables. For clarity in interpreting the factor analysis, the ARHQ scores were multiplied by -1 because unlike the rest of the tests, lower scores indicate *less* trouble reading.

Based on the elbow of the scree plot, interpretability, and amount of variance accounted for (See Figure 3), 5 factors collectively accounting for 59.5% of the variance were extracted (23.4%, 13.4%, 9.3%, 7.3%, and 6.1% respectively).

The first three factors were very similar to the factors found without the ARHQ included. The first factor is related to reading expertise, and the ARHQ dimension reflecting reading attitude and speed loads on this factor. The second factor reflects sublexical decoding and spelling skills. The third factor includes the Nelson-Denny accuracies and the Raven's scores, again with the Raven's speed being negatively correlated with the factor. No additional scores from the ARHQ loaded on this factor. The ARHQ introduced two additional factors. The fourth factor includes the learning and memory dimensions of the ARHQ, along with the reading disability factor, and the fifth factor includes only the "Casual Reading" dimension. This factor accounts for the least variability.

In summary, when the 5 dimensions extracted from the Adult Reading History Questionnaire were included in the factor analysis, the same 3 factors emerged as without

Figure 3: Scree Plot for Factor Analysis including ARHQ dimensions



(speed/reading experience, sublexical skill, and general accuracy/ability), along with an additional two dimensions of variability: learning and memory, and how much brief or casual reading one does.

2.4 DISCUSSION

The factor analyses of a database of adult reading assessments revealed the underlying patterns of variability in data measuring various skills relating to adult reading ability. One pattern that emerged as being quite consistent across analyses was that people who tend to like reading books and have a positive attitude about reading have more exposure to text and also tend to read faster, consistent with a view in which practice reading improves fluency and vice versa. This was true by both self-report on the ARHQ and in more objective test measures. In addition, when the ARHQ factors were included in a factor analysis with the

Table 5: Five factor PCA solution including ARHQ with varimax rotation

	Factor 1 Expertise	Factor 2 Sublexical	Factor 3 Accuracy	Factor 4 Learning/Mem	Factor 5 Casual Rdng
Vocabulary % Attempted	0.776	0.256	-0.151	0.075	0.002
Comprehension % Attempted	0.763	0.138	-0.223	0.101	0.044
ARHQ: Books/Attitude/Speed	0.655	-0.256	0.238	0.182	0.049
Author Recognition Test	0.56	0.408	0.259	-0.048	0.190
Real Word Test	-0.022	0.740	0.198	0.143	-0.081
Phonological Awareness	0.060	0.721	0.242	0.087	-0.103
Spelling	0.233	0.614	-0.081	0.129	0.108
Raven's % Attempted	0.236	0.078	-0.725	-0.088	-0.016
Raven's % Accuracy	-0.057	0.194	0.626	0.023	-0.101
Vocabulary % Accuracy	0.366	0.412	0.543	-0.041	0.102
Comprehension % Accuracy	0.246	0.337	0.483	0.000	-0.022
ARHQ: Reading Disability	0.067	0.081	0.130	0.741	-0.083
ARHQ: Spelling/Learning	0.253	0.192	0.075	0.713	-0.045
ARHQ: Memory	-0.068	0.038	-0.163	0.648	0.347
ARHQ: Casual Reading	0.120	-0.046	-0.032	0.041	0.918

Darkly shaded cells indicate a factor loading > .5

Light shading denotes highest loadings for questions that do not load heavily on any factor other assessments, the ARHQ expertise factor clustered with the objective measures of speed and experience. The speed component of the expertise factor was specific to reading speed; the number of questions answered in the time limit on the Raven's non-verbal intelligence test only loaded strongly (and negatively) on the accuracy factor.

This brings us to the next point: expert readers did not necessarily read more accurately, as accuracy on the comprehension, vocabulary, and Raven's Matrices tests (all of the timed tests) remained a separate factor across analyses. Accuracy scores were computed as the percent of attempted questions that were answered correctly. This means that a reader who answered only one question, but answered it correctly, would have the same accuracy score as a reader who answered all of the questions correctly. This method of scoring allowed us to separate speed from accuracy. Readers scoring highly on the accuracy factor, then, can be thought of as readers who progress at a pace that allows them to remain accurate. These readers have an accuracy focus. This is consistent with the negative loading of the Raven's speed measure. Because this test gets progressively more difficult, it is likely that readers who are able to answer more questions in the time limit are doing so at the expense of remaining accurate.

Sublexical skills, or those focused on orthography and phonology, were consistently clustered. There was no separation between the test that focused on orthographic knowledge (the spelling test) and those that focused on decoding and phonological awareness (the PhAT and the Real Word Test). Although this is the case for the full set of data, it is possible that orthographic, decoding, and phonological awareness skills would not be as tightly coupled in a subset of only less-skilled readers, as was tentatively reported by Perfetti and Hart (2002).

When the five factors extracted from the Adult Reading History Questionnaire were included in the analysis, results reproduced the expertise, sublexical, and accuracy focus factors, but could additionally account for variability in learning/memory, and "casual reading" (two additional factors). It is interesting that reading the internet, newspapers, and magazines is not highly correlated with reading attitude and reading speed. Reading books, on the other hand, is highly correlated with the speed/experience factor.

3.0 STUDY 2: RELATIONSHIP BETWEEN INDIVIDUAL DIFFERENCES AND EYE MOVEMENTS

3.1 INTRODUCTION

The factor analysis identified five main dimensions along which normal college readers vary, as measured by a wide battery of reading tests: (1) expertise (speed and experience) (2) sublexical skills, (3) accuracy focus, (4) learning/memory, and (5) amount of casual reading done. These sources of variability were defined using a data-driven approach, and thus represent true dimensions of variation in a large sample. The purpose of Study 2 is to understand whether scores along these dimensions are reliable predictors of eye movement patterns reflecting online reading processes.

We hypothesize that the first factor (expertise) will affect eye movements in the same way as word frequency because the variability is based on the same principles: frequent words have been encountered often, and those who do more reading have encountered words more often. Also, because this factor also included high weightings from both self-reported and measured reading speed, we expect to confirm that readers scoring highly along this dimension read words faster. Therefore, we predict shorter and fewer fixations, as well as more skipping, by those with more expertise.

It is generally thought that first-pass eye movement measures are influenced by all stages of word identification, whereas second-pass reading reflects meaning integration and other higher level reading processes (e.g. Kliegl, Grabner, Rolfs, & Engbert, 2004). Because sublexical skills should be important early in word identification, we expect to see an effect of the sublexical skill factor primarily in first pass measures (first fixation duration, probability of re-fixating the word, first pass refixation durations). It is not clear how scores along the accuracy focus factor are related to the fluency of reading processes. However, because this factor comprises accuracy on vocabulary, comprehension, and non-verbal reasoning tests, we might expect to see any effects on eye movements occurring in later measures (e.g re-fixations, second pass reading).

Neither the sublexical skill factor nor the accuracy focus factor have high weightings from the speed components of the given assessments. Instead, they reflect accuracies of the various assessments comprising the scores. Because of this, we may find that the primary effects of these differentiators appear in the form of interactions with the reading expertise factor. For example, a good comprehender who is fast should show a different pattern of eye movements than a good comprehender who is slow. And likewise, a fast reader who is also comprehending the text well may show a different pattern than a fast reader who is not a good comprehender. For this reason, we also examine the interactions between reading expertise and both the sublexical and accuracy focus scores to find out whether efficient reading behaviors depend on a combination of both practice and more specific sub-skills.

The last two factors, learning/memory and amount of casual reading explain less individual variability in the factor analysis and will thus likely be weaker predictors of reading behavior. However, we may see that those who learn more easily gain more from each reading experience. Thus, the main prediction for the learning/memory factor is an interaction with the reading experience factor. Perhaps the amount of reading experience is less important for those who have an easier time learning and remembering information. The last factor, amount of casual reading done, should in theory look no different than the first factor, expertise. However, because this factor has the smallest range of variability, we do not necessarily expect to be able to detect the effect of casual reading on eye movements.

Without attention to individual differences, many studies have focused on how the average reader interacts with words in particular text structures or with particular lexical properties. It would be easy to think of lexical factors as being intrinsic to each word, but few of the many commonly examined lexical properties actually are. Word length or number of syllables are good examples of such fixed, context-independent lexical properties. However, neighborhood size, for example, is a lexical property contingent on the rest of the reader's lexicon. Any effects of neighborhood size would be premised on a reader having knowledge of each of the enumerated neighbors, a problem in neighborhood measurements that has been proposed to affect experimental results (Andrews, 1997). Word frequency or the mean frequency of a word's neighbors are lexical properties that depend on how often words appear in text, and are not inherent to the word. Again, effects of these properties are premised on a reader having encountered the words proportionally to their occurrences in a particular corpus. Not only do these properties depend on a reader's experience or knowledge, but also on the skills that determine how effectively a reader can learn from experience or use their knowledge.

Therefore, in addition to predicting main effects of and interactions between individual difference scores, we expect to see interactions between particular individual differences and related lexical factors. Specifically, we expect to see interactions between frequency-related lexical factors and reading expertise. Experienced readers could either be less sensitive to frequency effects because of their many encounters with less-frequent words, or they could be more sensitive to frequency effects because their actual experience with words is closer to estimated frequencies than it is for those who do less reading. Alternatively, reading expertise and word frequency may have simple additive effects. We also expect to find an interaction between word frequency and sublexical skill. If readers rely more heavily on sublexical skills for reading less-frequent words than they do for more-frequent words, we would expect to see a larger relationship between fixations and sublexical skill for less-frequent words than for more-frequent words (similar to behavioral effects in children discussed in Juel, 1980). We may also find a relationship between sublexical skill and bigram frequency. If words with low bigram frequencies are more difficult to decode, we may see a larger effect of sublexical skill for such words. Alternatively, if high bigram frequency words are more confusable, we may see a larger effect of sublexical skill for the high bigram frequency words.

The results of Study 2 supplement our understanding of word-based and text-based influences on reading behavior with an understanding of individual skill-based and experience-based influences. Results highlight that individual skills and experience help explain variability in patterns of reading above and beyond lexical-level factors, and results also show that, in fact, the effects of lexical properties are mediated by individual differences. In addition, because there are a number of measurable properties of eye movements, we can achieve a more fine-tuned understanding of which particular online reading behaviors are predicted by each skill, or when during reading each skill is important.

3.2 METHODS

3.2.1 Procedure

Sixty-eight paragraphs were obtained from the internet with some slight modifications made for clarity out of context. They ranged from 80 to 119 words per paragraph, with a mean of 95 words per paragraph. A norming study was conducted to ensure that the paragraphs were of average difficulty and were fairly homogenous in their difficulty level. Each paragraph was rated for difficulty by 33 native English speakers who received course credit. Participants read the paragraphs on paper, and used a 7-point scale (ranging from 1 to 7) to rate the difficulty, on which 1 was "very easy", 4 was "average," and 7 was "very difficult." These were defined as follows:

Very easy: A very easy passage is well below your level of reading ability. It can be completely understood with minimal effort.

Average: An average passage matches your normal level of reading difficulty. You can read the passage at a normal pace and understand the text.

Very difficult: A very difficult passage is beyond your normal level of reading difficulty. You might need to spend a long time on the passage, re-read the passage, or read effortfully to understand the text. You may feel you need more surrounding context to understand the passage.

The mean paragraph rating was 3.24 (SD = .59), or a little below average difficulty, with a minimum average paragraph rating of 2.21 and a maximum of 4.54.

A group of 35 native English speakers read each of these 68 paragraphs while their eye movements were monitored. Paragraphs were presented in a randomized order. A True/False comprehension question with feedback was given following each of the paragraphs to ensure that participants stayed focused on reading the passages for meaning.

Eye movements were monitored from the right eye using an Eyelink 1000 eyetracker with a sampling rate of 1000 Hz. We used a standard 9 point full-screen calibration before each participant began reading. A center point only calibration was used between each trial, and a full 9 point calibration was re-conducted as necessary throughout the experiment. Participants were instructed to read each paragraph once for meaning, and to press a mouse button when they were finished reading. If a True/False question appeared after reading the paragraph, participants would use the left mouse button to indicate "true," and the right mouse button to indicate "false," corresponding to the side of the screen on which the answer choices "true" and "false" were printed. The total duration of the experiment depended on participant's speed, but was approximately one hour.

3.2.2 Dependent Variables

Vertical drift was corrected using the Eyelink Data Viewer software, and trials in which the calibration was appreciably off-target were removed. In addition, any fixations less than 50 ms within .5 character spaces of another fixation were merged. For words that were not initially skipped, the following fixation durations were computed:

First fixation duration is the duration of the first fixation on a word. *Refixation duration* is the cumulative duration of any additional fixations made on the first pass reading of the word (before the reader has moved their eyes off of the word). *Gaze Duration* is the cumulative duration of all first pass fixations (the sum of the first fixation duration and refixation durations). *Re-reading duration* is the cumulative duration of any fixations made on the word after the eyes have moved off of the word (forward or backwards). *Total Viewing Time* is the sum of all fixations on the word. Only cases for which refixations or re-reading occured were included in the refixation and re-reading durations. Log transforms were applied to each of the duration measures to normalize the distributions.

In addition to measuring the durations of fixations, we also predicted probabilities that a word would be initially skipped, the probability that a word would be refixated on the first pass (given that it was not skipped), and the probability that a word would be re-read after the first pass reading. Additionally, for words that were initially skipped, the probability of looking back at the word vs. never viewing the word was measured.

3.2.3 Independent Variables

The individual difference variables of interest were the five factors derived from the factor analysis and the operation span working memory score. Although the factor scores are centered and follow a normal distribution over the whole population of subjects included in the factor analysis, these scores were re-centered for the subset of participants who also participated in Study 2 so that the zero point represented the experimental sample mean. Trial number was also centered and included as an independent variable.

Lexical factors were also included in the model. The following ten lexical properties were obtained from the English Lexicon Project database for 1,167 of the words appearing in the paragraphs: number of syllables, number of phonemes, length, orthographic neighborhood size, phonological neighborhood size, log frequency of the orthographic neighbors, log frequency of the phonological neighbors, number of morphemes, bigram frequency by position (we computed the mean bigram frequency by position by dividing by the number of bigrams in the word), and log HAL frequency (Balota et al., 2007).

Many of these variables are highly correlated (such as number of syllables and number of phonemes) and are thus not suitable for use together as predictors. As a compromise between eliminating a subset of the variables and including them all despite the interpretation and replicability problems associated with collinearity, we conducted a PCA with varimax rotation to compute meaningful, orthogonal dimensions along which words vary. Blending the word properties in this way was acceptable because it was not the goal of the experiment to parcel out the individual influences of covarying lexical properties.

3.2.3.1 Factor Analysis of Lexical Attributes A PCA with Varimax rotation was performed on the ten lexical properties listed above. 89.75% of the variance could be accounted for by reducing the ten lexical properties to six factors, and these 6 factors were highly interpretable. The first factor is a word length factor that included the number of syllables, number of phonemes, and number of letters. The second factor reflected the number of neighbors, with high loadings from both the number of orthographic neighbors (differing by one letter) and the number of phonological neighbors (differing by one phoneme). The third factor was a neighborhood frequency factor and had high loadings from both the orthographic and phonological neighborhood frequency measures (reflecting how frequent a word's neighbors are on average). The final three factors had high loadings from only one measure each: number of morphemes, mean bigram frequency by position, and log HAL frequency, respectively. These 6 orthogonal factors (corresponding to length, number of neighbors, neighborhood frequency, number of morphemes, bigram frequency, and frequency) were used as independent variables predicting eye movements.

3.2.3.2 Interactions Across the various models, we tested for the following predicted interactions, as discussed in the chapter introduction:

- expertise x sublexical skill
- expertise x accuracy focus
- expertise x word frequency
- expertise x neighborhood frequency
- expertise x bigram frequency
- sublexical skill x word frequency
- sublexical skill x bigram frequency

3.2.4 Data Analysis

Data for reading times were analyzed using Linear Mixed Effects Regression (LMER) analyses (Baayen, 2008; Baayen, Davidson, & Bates, 2008) . Similar to a linear regression analysis, the LMER produces a linear model in which the dependent variable is predicted by the sum of each independent variable multiplied by a coefficient. Independent variables are considered to be reliable predictors of the dependent variable if they have coefficients that are determined to be significantly different than zero. Because our independent variables are scores along dimensions produced by a factor analysis, and because our dependent variables are often transformed so that they can be predicted by a linear model, it can be difficult to interpret the meaning of the values of the coefficients. For example, a coefficient of .05 for the effect of word length on log total viewing time means that for every increase of .05 standard deviations along the word length factor, there is a corresponding increase of 1 in the log total viewing time.

The benefit of the LMER above a linear regression is that the LMER allowed us to include subjects, words, and paragraphs as random factors in a single analysis. For all models, only the intercept was allowed to vary by the random factors, with the exception that the slope of trial number was allowed to vary by subject to account for any practice or fatigue effects (with no random correlation term between the intercept and trial number slope). Slopes were allowed to vary only via fixed effect interactions of interest (as specified above) – further allowing the slopes to vary by random effects would probably overfit the model, as evidenced by the failure of these models to converge in a reasonable number of iterations.

The models were fit using restricted maximum likelihood. *P*-values were obtained based on highest posterior density confidence intervals computed using Markov chain Monte Carlo (MCMC) sampling with 10,000 iterations (see Baayen, 2008, p.270). This avoids anti-conservative *p*-values that can arise from use of the *t*-statistic with the upper bound of degrees of freedom.

Probabilities of skipping, refixating, and re-reading (binary coded variables) were analyzed with mixed-effects logistic regression which uses a logit transformation of the binary dependent variable. Again, intercepts for these models were allowed to vary by the subjects, words, and paragraphs as random factors. The slope for trial number was allowed to vary by subject to account for fatigue or practice effects. *P*-values for this set of models were computed using the Wald Z statistic.

3.3 RESULTS

All estimated model coefficients are summarized in Table 3.3.

Table 6: Study 2 Model Coefficient Estimates

		Initially Viewed Words							Skipped Words	
		TVT	P(Refix)	P(Reread)	FFD	Refix	GD	Reread	P(Skip)	P(Reread)
	Intercept	5.454***	-2.571***	-1.811***	5.267***	5.101***	5.322***	5.372***	-0.076	-1.029***
	Trial Number	-0.001*	-0.003**	-0.005***	0.000	-0.001	0.000	0.000	0.001	-0.005***
Lexical Factors	Length	0.054***	0.252***	0.113***	0.016***	0.020***	0.040***	0.040***	-.3.460***	0.268***
	Neighbors	-0.031***	-0.188***	-0.079***	-0.008**	-0.017**	-0.022***	-0.014*	0.216***	-0.182***
	Neighbor Freq	-0.022***	-0.118***	-0.072***	-0.007**	-0.007	-0.014***	-0.004	0.189***	-0.099**
	Morphemes	0.035***	0.209***	0.097***	0.004	0.009	0.023***	0.014**	-0.250***	0.182***
	Bigram Freq	0.017***	0.122***	0.047*	0.002	0.013*	0.011***	0.006	-0.094***	0.096***
	Frequency	-0.041***	-0.177***	-0.059**	-0.018***	-0.037***	-0.032***	-0.016**	0.225***	-0.216***
Individual Differences	Expertise	-0.062**	-0.246***	-0.241**	-0.025	-0.049*	-0.037*	-0.027	0.212***	-0.337***
	Sublexical	-0.046	-0.214**	-0.131	-0.029	-0.026	-0.038	-0.041	0.157*	-0.154
	Accuracy Focus	0.060*	0.097	0.376***	0.020	0.036	0.024	0.034	0.122	0.269*
	Learning/Memory	-0.022	0.010	-0.095	-0.010	-0.018	-0.011	-0.030	0.052	-0.119
	Casual Reading	0.001	-0.064	-0.014	0.004	0.005	0.001	-0.008	0.080	-0.004
Interactions	Expertise*Sublexical	-0.040	-0.144	-0.050	-0.015	-0.007	-0.021	-0.037	-0.071	-0.200
	Expertise*Accuracy Focus	0.017	0.181*	0.217*	-0.009	0.017	0.002	-0.019	-0.030	0.226
	Neighbor Freq*Expertise	0.000	-0.015	-0.027*	-0.002	0.010	0.000	0.003	-0.009	0.018
	Frequency*Expertise	0.004**	-0.014	-0.005	-0.001	0.000	0.002	0.003	-0.003	-0.002
	Bigram Freq*Expertise	-0.002	0.005	0.013	-0.002	-0.004	-0.003*	-0.001	0.007	0.006
	Frequency*Sublexical	0.000	-0.002	0.001	-0.003*	-0.004	-0.001	0.003	0.003	-0.028***
	Bigram Freq* Sublexical	0.000	0.012	-0.009	0.001	0.004	0.001	-0.006	0.009	-0.007

*$p < .05$; **$p < .01$; ***$p < .001$

3.3.1 Lexical Factors

3.3.1.1 Initially Viewed Words The total viewing time was affected by each of the lexical factors. As expected, shorter, more frequent words with many frequent neighbors and fewer morphemes were read more quickly. Of note is the relationship between the bigram frequency factor and total viewing time. More time was spent reading words with higher bigram frequencies. The bigram frequency factor here may be more indicative of how informative the letter sequences are, with lower bigram frequencies being more informative about the word identity, as has been reported to be the case especially in low frequency words (Broadbent & Gregory, 1968).

Total viewing times depend on the combination of fixation durations and the probabilities of refixating. Each of the lexical factors contributed reliably to the both the probability of refixating the word on the first pass and the probability of re-reading the word once it had been left. Thus, a clear way that these lexical factors have an effect on viewing times is by increasing or reducing the number of fixations on a word.

Another way that lexical factors can contribute to longer total viewing times is by modulating the durations of fixations, including the first fixation duration, refixation durations, and re-reading durations. Results show that there are some differences in which fixation durations are affected by each factor. Word length, neighborhood size, and word frequency affect fixation durations across the board. The frequency of the neighbors affects the first fixation duration, but does not reliably affect re-fixation durations or re-reading durations. Bigram frequency affects refixation durations, but not first fixation durations or re-reading durations. The number of morphemes in a word affects re-reading durations, but not first pass fixation durations.

Each of the lexical factors affect gaze duration, so the combination of effects on the first fixation durations, the probability of refixating, and the refixation durations (all the first pass measures) indicate that all of these lexical factors are important in modeling first pass reading behavior.

3.3.1.2 Skipping

All of the lexical factors also helped predict whether a word would be skipped. Again, the directions of these effects were as expected, and again, it was the lower bigram frequency words that were more likely to be skipped. For words that were initially skipped, all of the lexical factors contributed to predicting whether the word would then be later reviewed (vs. skipped completely). More difficult words (those with longer lengths, lower frequencies, more morphemes, higher bigram frequencies, etc.) were more likely to be re-read after being initially skipped.

3.3.2 Individual Differences

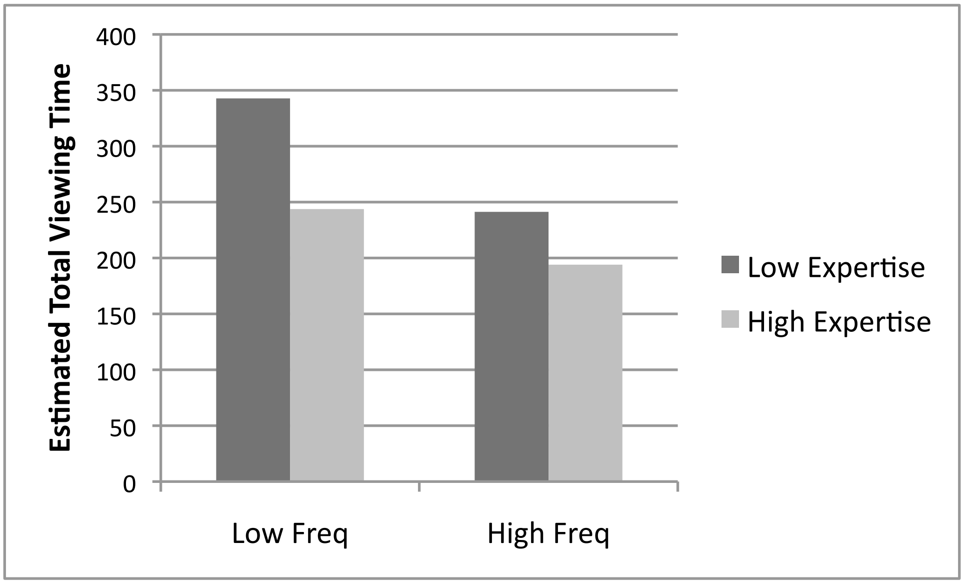
3.3.2.1 Initially Viewed Words

Total viewing times for non-skipped words were affected by two of the five individual difference factors: the expertise factor and the accuracy factor. There was also a trend towards sublexical skills reducing total viewing times ($p < .08$). High scores on the expertise factor were not surprisingly associated with shorter total viewing times. Higher scores on the accuracy focus factor, which reflected accuracies on the Nelson-Denny tests as well as scores on the non-verbal intelligence test, were associated with longer total viewing times. In addition, the word frequency effect was modulated by reading expertise; more-experienced readers showed less of a frequency effect (see Figure 4). This interaction was driven by the expertise advantage being greater for less-frequent words. There was a trend towards higher sublexical skills decreasing total viewing times ($p < .08$).

These effects on total viewing times come from individual differences in fixation durations and probabilities of refixating. Those with more expertise and better sublexical skills have a lower probability of refixating words after the first fixation. There is also an expertise x accuracy focus interaction in which the expertise effect is stronger for the low accuracy focus readers. Lower accuracy focus/lower expertise readers are the most likely to refixate, but low accuracy/high expertise readers are least likely to refixate. The latter group may represent careless readers who prefer to read the text quickly at the expense of accuracy.

The probability of re-reading words after moving off of them shows a similar pattern. Those with higher expertise are less likely to re-read. Those more focused on accuracy are more likely to re-read. There is an interaction between the two groups that likely drives most

Figure 4: Interaction Between Reader Expertise and Word Frequency on Total Viewing Time



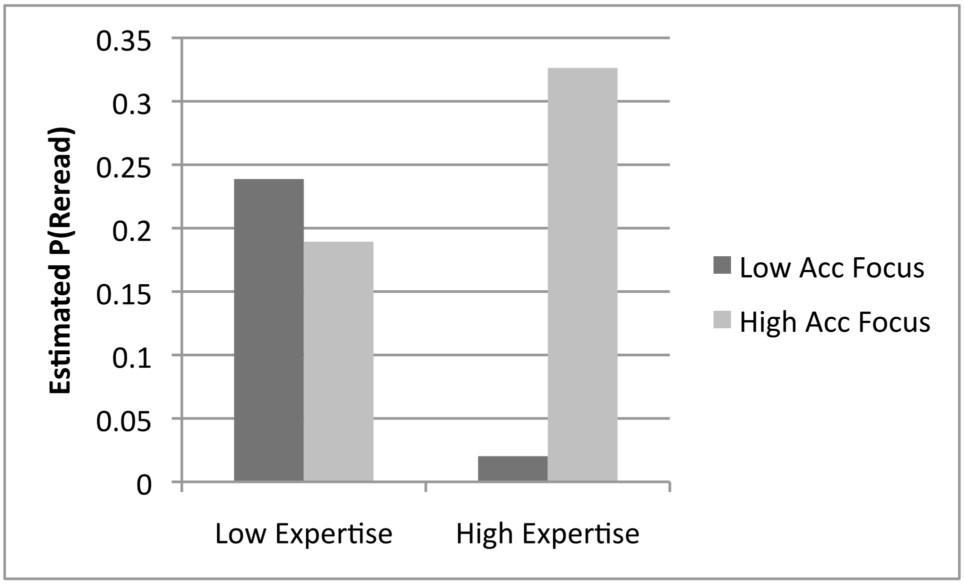
Graph of model predictions for the bounding data values illustrating the interaction between reader expertise and word frequency.

Predictions are back-transformed from log fixation durations for display in msec.

of the main effect, which shows that it is primarily the low accuracy/high expertise (fast inaccurate) readers who show a low probability of re-reading (see Figure 5). There is also an interaction between neighborhood frequency and expertise whereby the experienced readers show a greater advantage (less probability of re-reading) for high neighborhood frequency words, and thus a greater overall neighborhood frequency effect.

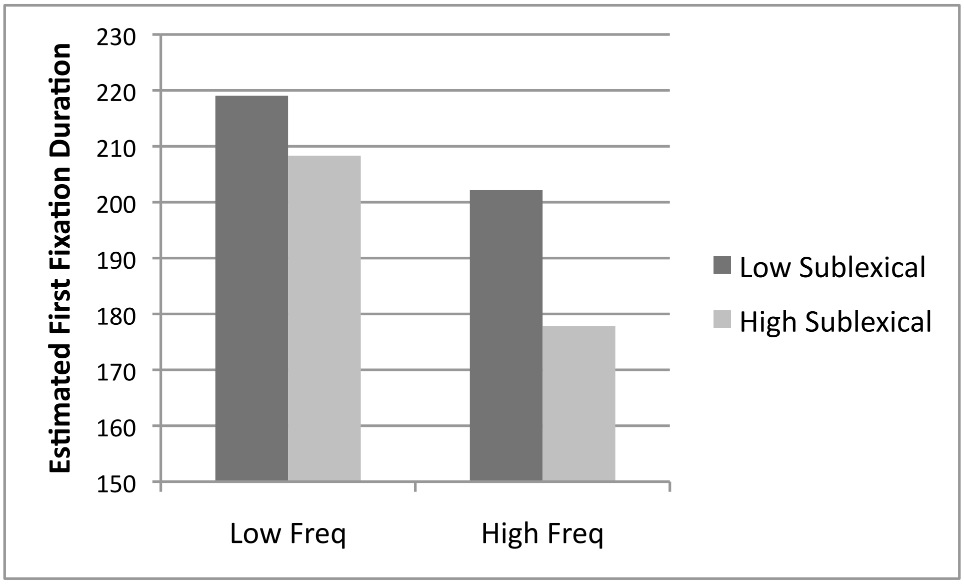
In addition to the individual differences in probabilities of refixating and re-reading, there were also individual differences in the durations of fixations. For the first fixation durations, those with good sublexical skills showed a greater advantage for high-frequency words, and thus a larger frequency effect (see Figure 6). More-experienced readers had shorter refixation durations, and shorter total gaze durations, especially for high bigram frequency words. Whereas there was not a large bigram frequency effect for experienced readers, the less-experienced readers had longer gaze durations for high bigram frequency words. There were no significant individual differences in re-reading durations.

Figure 5: Interaction Between Reader Expertise and Reader Accuracy Focus on Probability of Rereading



Graph of model predictions for the bounding data values illustrating the interaction between reader expertise and accuracy focus.

Figure 6: Interaction Between Reader Sublexical Skill and Word Frequency on First Fixation Durations



Graph of model predictions for the bounding data values illustrating the interaction between sublexical skill and word frequency.

Predictions are back-transformed from log fixation durations for display in msec.

3.3.2.2 Skipping Readers with more expertise and better sublexical skills were more likely to skip words. Once a word had been skipped, those with more expertise were also less likely to go back and read the word. Those with lower accuracy scores were less likely to go back and re-read the skipped words. There was also an interaction between sublexical skill and word frequency in predicting the likelihood of re-reading skipped words: Having higher sublexical skill only benefitted readers for higher frequency words – the high sublexical group were less likely to re-read the more-frequent words.

3.4 DISCUSSION

The results of this experiment showed that in addition to the contributions of lexical factors to reading durations, individual differences play a significant role. The major findings were

(1) widespread main effects of reading expertise (a combination of experience, attitude, and speed), (2) interactions between reading expertise and frequency-related measures, (3) early effects of sublexical skills, (4) later effects of the accuracy focus factor, and (5) interactions between expertise and accuracy.

The factor analysis reported in Chapter 2 showed that variability among normal college readers on a variety of reading tasks tended to fall along a few dimensions. The largest source of variability was in the readers' attitudes about reading, how much reading they tended to do, and the speed of reading, which all covaried. This was an important factor in predicting reading behavior. People with more experience reading and a positive attitude about reading exhibited faster reading behavior as a result of skipping more words, refixating less often and doing less re-reading, and having shorter refixation durations when they did refixate.

The amount of experience that readers had also modulated frequency-related effects. The nature of the interactions varied depending on the nature of the frequency effect. For word frequency, the expertise advantage for total viewing times was particularly evident for lower frequency words. This result suggests that increasing numbers of encounters with any particular word improve reading efficiency with diminishing returns. In other words, extra encounters with less-frequent words improve reading more than extra encounters with already frequently encountered words.

Reading expertise interacted similarly with bigram frequency in predicting gaze duration. The bigram frequency effect resembled the word frequency effect in that there was less of a bigram frequency effect for more-experienced readers. The words that less-experienced readers fixated longer (high bigram frequency words) were not fixated longer by experienced readers. This is a bit counter-intuitive. If similarity to other letter sequences provides interference, it might be predicted that those who do more reading would experience greater levels of interference from high bigram frequency words. One possibility is that they experience the inhibitory effect more quickly, during parafoveal preview, which we did not examine in this experiment (White, 2008). Alternatively, more experiences with words may cause higher levels of specificity in word representations, making experienced readers less prone to confusion from similar letter strings (Andrews & Hersch, 2010; Castles et al., 2007).

Unlike the word frequency effect and bigram frequency effect, which were diminished for experienced readers, the effect of neighborhood frequency on the probability of re-reading was enhanced for experienced readers. The facilitation resulting from a word having high frequency neighbors was enhanced for those with more reading experience. It has been suggested that large neighborhood sizes and high frequency neighbors especially facilitate the recognition of low frequency words by boosting activation via form-related words (Andrews, 1989, 1992; Sears, Hino, & Lupker, 1995). Thus, more-experienced readers may benefit because they have had more exposure to similar neighbors, increasing activation of the target word via activation from neighbors, and decreasing the probability of re-reading the target word.

Results also showed that readers with good sublexical skills had shorter first fixation durations for more-frequent words, and had a lower probability of refixating words on the first pass. They were also more likely to skip words, and less likely to then go back and read more-frequent words. These findings demonstrate a continued effect of sublexical skills even in normal college readers. Contrary to predictions, however, sublexical skill had more of an effect on the processing speed of frequent words than infrequent words. This may be evidence of the idea that additional exposure to words causes them to become "unitized," a qualitative shift that results in more integrated word representations and less reliance on the processing of component features, and that unitization continues throughout adulthood (Spieler & Balota, 2000). More unitized words result in larger frequency effects, and Spieler and Balota found that older readers (with more exposure to words) showed these larger frequency effects. Interestingly, the finding that those with better sublexical skills show a larger word frequency effect could result from a mechanism whereby good sublexical skills help increase the rate of unitization with more exposures to words. A more integrated word form representation would speed early recognition or familiarity.

The accuracy focus factor was also related to total reading times. This factor included accuracies on the Nelson-Denny vocabulary and comprehension test, as well as accuracy on the Raven's matrices non-verbal test, and was negatively correlated with the number of questions answered on the Raven's. Participants scoring highly on this factor, then, will make sure to progress through tests at a rate that preserves accuracy. Thought of this way, the reading results make sense. Readers who focus less on accuracy generally have shorter total viewing times for words. This is partially due to less re-reading, both for skipped and not skipped words. The probability of re-reading, as well as the probability of re-fixating on the first pass, interact with the expertise measure such that it is specifically readers who are both fast and inaccurate who are less likely to re-read or re-fixate. These fast and inaccurate readers may be thought of as careless readers. This is possibly a reflection of the readers' behavior in the experimental context rather than their underlying skill.

No effects of the last two individual difference factors – a learning/memory factor and a casual reading factor – were found. There are probably several reasons for this. First, these two factors explained the least variability among normal college readers, making the factors less useful as possible predictors. Second, the experimental task was a casual reading task in which learning and memory may not have been as important as they are in other reading contexts. Lastly, it is interesting to note that despite the experimental task demands, having more experience in similar reading conditions (casual reading conditions) did not improve reading efficiency. Instead, the major effects were found for readers who have more experience with book reading.

In summary, Study 2 showed that individual differences play a significant role in reading behavior. Readers with good sublexical skills are facilitated in early measures, especially for more-frequent words, and there was a trend towards this effect reducing total viewing times. Those with more reading experience and who are generally faster readers had shorter total viewing times, a result of smaller probabilities of re-fixating, shorter re-fixation times, and less probability of re-reading. Those with lower accuracies also had shorter total viewing times. This was due to lower probabilities of refixating and re-reading, especially for the inaccurate readers who were also fast and experienced and might be considered careless.

Only sublexical skill modulates the reader's first fixation duration. This result stresses the importance of good sublexical skills in the earliest stages of word recognition. Interestingly, the facilitation may not result from fast decoding, but instead from a qualitative shift, supported by good sublexical skills, to more unified word representations, especially for frequent words. Thus, the size of the frequency effect on first fixation duration may be a useful measure for how integrated a reader's word representations are. Reading attitude, experience, and speed became important for faster reading subsequent to the first fixation. However, faster was not always better, as those readers who did not refixate or re-read as often may have had reduced comprehension in this reading task, as they did in the battery of tests given prior to the eye-tracking portion of the experiment.

The pattern of findings in this experiment suggest that there may be at least three important elements of individual differences in reading behavior: skills, knowledge, and strategy. Skills include overall reading speed and sublexical skills. Knowledge includes the quality and nature of lexical representations that result from the combination of the readers' skills and experience. Strategy includes whether a skilled reader may be reading carelessly, and reflects the readers' goals.

Main effects of the expertise factor may reflect the quality of lexical representations, but may also simply reflect practice and skill with the low-level motor attributes of reading. The pattern of interactions of this factor and the sublexical skill factor with frequency-related word attributes is more strongly suggestive that the quality of word representations and knowledge of specific words plays a significant role in reading behavior. However, reading behavior can also resemble more efficient reading for participants who are actually not taking the time to accurately comprehend the text. Thus, strategy or reader goals must be taken into account when interpreting individual differences in reading behavior.

4.0 STUDY 3: EFFECTS OF PARTIAL WORD KNOWLEDGE ON EYE MOVEMENTS (TRAINING PARADIGM)

4.1 INTRODUCTION

The collective results from a factor analysis of individual differences in reading skills and an analysis of how those skills predict reading behavior indirectly support the idea that word knowledge plays an important role in reading behavior and in accounting for individual differences in reading behavior, in addition to skill and strategy. The purpose of Study 3 is to more directly examine the link between word knowledge and reading behavior at the single word level by controlling word knowledge experimentally.

There are multiple components to knowledge of a word, including orthographic, phonological, and meaning components, and there is consensus that each of these components is accessed during reading, as often evidenced through studies that manipulate form and meaning features in priming and lexical decision tasks (Lee, Rayner, & Pollatsek, 1999; Perfetti, Bell, & Delaney, 1988; Perfetti & Bell, 1991; Rayner, Sereno, Lesch, & Pollatsek, 1995). With high levels of word knowledge in skilled readers, representations of these components should be stable and tightly linked. It is possible that word frequency or familiarity effects are due at least in part to reduced or strengthened word representations across all components. Generally, words that are often seen are also often heard and have well-known meanings. Barry, Morrison, and Ellis (1199), for example, report correlations ranging from .75 to .89 between log spoken frequencies and the log frequencies of three written frequency measures. Familiarity ratings were also highly correlated with all of the frequency measures, with correlation coefficients ranging from .47 to .61. Because form information (especially orthography in the case of reading) is more quickly accessible, early decisions about when to move the eyes forward and whether to refixate the same word may be based on form familiarity alone, with the probable outcome that the form familiarity is a good indicator of how easy the meaning will be to access. However, even with "known" words, there may be weaknesses in one or more of the components of representation. Perhaps a word has been heard many times but rarely encountered in text or that a reader isn't quite sure how to pronounce a word they see in text. Alternatively, a reader may recognize a word but be somewhat unsure of its meaning, with context triggering a vague recollection.

The purpose of Study 3 is to determine how the strength of each component of word knowledge is manifested in the eye movement record. This is achieved by controlling the amount of exposure each participant has to the visual, auditory, and definition information for rare, previously unknown words in a vocabulary training paradigm. Of course we can not completely control the quality of representations for each of the components for several reasons. One is that there will be individual differences in how much each person learns from each exposure. Also, even without explicit visual or auditory presentations during training, readers (especially those with good sublexical skills) should be able to assemble the missing information from the form information they are given via letter-to-sound correspondences. Nevertheless, we make the assumption that explicit training of a particular component results in stronger knowledge of that component, as do more exposures during training.

We reasoned that by observing eye movements when participants read sentences containing the trained words, we could determine how different types of low quality representations might differentially affect reading behavior. This will result in a more direct understanding of the role each component of word knowledge has in reading behavior. We also examine how individual differences interact with the training to produce distinct patterns of reading.

Specific hypotheses are (1) orthographic training will have the earliest effects because the reader will predict how long word identification will take based on visual familiarity and plan subsequent eye movements accordingly. (2) Meaning knowledge will become important in later measures when initial form-based assumptions about how well-known the word is prove to be correct or incorrect. (3) Explicit phonological training will be most helpful for readers with low sublexical skills, because good decoders will not get much extra benefit from hearing the word pronounced (based on J. R. Nelson et al., 2005; Share, 1995). (4) Readers with more expertise will benefit more from training because of more extensive baseline knowledge with which they can integrate the new information (Perfetti et al., 2005; Balass et al., 2010) .

4.2 METHODS

4.2.1 Participants

35 native English speakers who had taken the battery of reading tests or were willing to take the battery of reading tests were recruited to participate for payment.

4.2.2 Materials

180 rare words were chosen from a set of normed words. The words were rated for familiarity and knowledge of the word's meaning by 30 participants receiving course credit for their participation. Each of these ratings was completed using a 7-point scale ranging from "unfamiliar" to "familiar" for the familiarity rating, and from "not at all" to "completely" for how well the definition was known. Words were chosen to have low familiarity and meaning ratings. The mean familiarity was 2.18 (ranging from 1.33 - 3.00), and the mean meaningfulness was 1.78 (ranging from 1.17 - 2.96). Words with extremely low familiarity and meaning scores were not chosen because they tended to appear very unusual and non-wordlike.

All words were either 7 or 8 letters long. Longer words were chosen to reduce the chance of skipping during reading. Words were divided into 18 lists of 10 words each, balanced for mean word length and part of speech. These words lists were assigned to a 6 (training condition) x 3 (number of exposures) training design. The six training conditions were the following combinations of orthography (O), phonology (P), and meaning (M): O, P, OP, OM, PM, and OPM (every combination except for meaning only). Completely untrained words were not included, because a pilot study showed that participants exhibited highly inflated looking times when reading the untrained rare words in context, re-fixating the words as many as 20 times. Participants may have been trying to recall whether they had learned the word, rather than reading the word normally. Thus, trained conditions were always significantly faster than untrained conditions, but not necessarily in a meaningful way. We hoped to increase the degree to which participants would read text normally rather than stare at target words by leaving out the untrained condition. Words in each training condition appeared one, three, or five times during the training. A partial latin square design was used to rotate the word lists through the experimental conditions, resulting in six versions of the experiment.

4.2.3 Training Portion

Presentation of experimental words was randomized (conditions were not blocked). Conditions including the orthography showed the word written in lowercase letters. Conditions including phonology showed a speaker icon which then disappeared when a recording of the word's pronunciation was played over the computer speakers. Conditions including meaning showed a brief definition for the word. To minimize variability in the amount of exposure each participant had to the words, training was not self-paced. Instead, words in each condition were presented for enough time for participants to be able to briefly study the information. The conditions with more information to attend to were presented for longer durations. The amount of time the display remained on screen for each condition was as follows: O: 3000 ms, P: 4000 ms (1000 ms before the sound played + 3000 ms of a blank screen after), OP: 5000 ms (2000 ms before the sound played + 3000 ms after), OM: 7000 ms, PM: 7000 ms (2000 ms before the sound played + 5000 ms after), OPM: 9000 ms (2000 ms before the sound played + 5000 ms after).

To ensure that participants were paying attention to the training, probe questions followed a random 10 % of trials. The questions were appropriately tailored to the training condition. Words that were trained with multiple components of knowledge were followed by a question about any one of the components given. Questions included:

- Orthography: What letter did the previous word begin with?
- Orthography: What letter did the previous word end with?
- Phonology: How many syllables were there in the word?

- Phonology: Did the word start with a vowel or consonant sound?
- Meaning: Is the meaning a thing, action, or description? (participants were instructed that these choices correspond to nouns, verbs, and adjectives)

Prior to training, participants were given seven practice words with the experimenter to become familiar with the pacing and question probes. One of the seven practice words was presented twice, and one of them was presented three times, so that participants would learn that words would sometimes be presented multiple times (and this was also explicitly described by the experimenter). Three of the example words were also followed by example probe questions.

Participants were encouraged to take breaks as needed before proceeding forward from probe questions. The total time spent in training was approximately 1.5 hours.

4.2.4 Eyetracking Portion

Following training, participants read one sentence per trained word while their eye movements were monitored. Sentences did not span more than one line of text. The trained words were never the first or last word in a sentence, nor did they appear before a comma or other punctuation. Each sentence was followed by a True/False comprehension question to ensure that participants were motivated to read for meaning. The experimenter was in the room with the participant for the whole session. Calibration procedures and eye-tracking details were the same as for Study 2.

4.2.5 Dependent Variables

Dependent variables were computed in the same manner as in Study 2 with the exception that skipping measures were not computed for this data set. This is because by choosing longer target words, we ensured that few target words were skipped.

4.2.6 Independent Variables

Individual difference variables of interest were the five factors derived from the factor analysis, minus the casual reading factor. Because this factor explained the least variability, showed no effects in Study 2, and was not hypothesized to predict effects in Study 3, it was removed from the analysis. The Learning/Memory factor was included in the analysis despite its lack of predictive power in Study 2 because Study 3 is a learning task, and may rely more heavily on this skill. Scores were re-centered for the subset of participants who participated in Study 3 by subtracting the sample mean so that the zero point represented the experimental sample mean. Lexical factors were not included in these models. This is because all of the words were of equal length (7-8 letters) and were very low frequency, unknown words.

Trial number was included to account for fatigue or practice effects. The number of exposures was also included as a numeric variable.

Training conditions were clustered for comparison in order to isolate effects due to orthographic, phonological, and meaning training. The comparisons highlight how the addition of a single element of training affects the reading outcome. (The lack of a meaning-only condition meant that the design is not completely balanced.) These are the compared conditions for each element of training:

- Orthography: P & PM vs. $\underline{\text{O}}$P & $\underline{\text{O}}$PM (addition of visual training)
- Phonology: O & OM vs. O$\underline{\text{P}}$ & O$\underline{\text{P}}$M (addition of pronunciation training)
- Meaning: O & P & OP vs. O$\underline{\text{M}}$ & P$\underline{\text{M}}$ & OP$\underline{\text{M}}$ (addition of definition training)

This characterization means that the comparison case for both the addition of orthographic and phonological information is the same (the OP and OPM conditions). For analysis purposes, this meant that a single independent variable with three levels (OP, no P, no O) was used to examine orthography and phonology training, with the combined OP and OPM training cases used as the baseline level for comparison (because it is the comparison level for both O and P training). A separate independent variable was used to examine meaning training.

4.2.6.1 Interactions Across the various models, we tested for the following interactions:

- expertise x orthography/phonology training
- expertise x meaning training
- expertise x number of exposures
- sublexical skill x orthography/phonology training
- sublexical skill x meaning training
- sublexical skill x number of exposures
- learning/memory x number of exposures
- orthography/phonology training x number of exposures
- meaning training x number of exposures

These interactions tested the hypotheses discussed in the chapter introduction that there will be interactions between individual skills and the training conditions, with more skilled or knowledgeable readers learning more from the training. Interactions between the training conditions and the number of exposures should reveal training effects that might become more pronounced with more exposures.

4.2.7 Data Analysis

Data for reading times were analyzed using Linear Mixed Effects Regression (LMER) analyses (Baayen, 2008; Baayen et al., 2008). This allowed us to include subjects and words as random factors in a single analysis. For all models, only the intercept was allowed to vary by the random factors, with the exception that the slope of trial number was allowed to vary by subject to account for any practice or fatigue effects (with no random correlation term between the intercept and trial number slope).

The models were fit using restricted maximum likelihood. P-values were obtained based on highest posterior density confidence intervals computed using Markov chain Monte Carlo (MCMC) sampling with 10,000 iterations (see Baayen, 2008, p.270). This avoids anti-conservative p-values which can arise from use of the t-statistic with the upper bound of degrees of freedom.

Probabilities of refixating and re-reading (binary coded variables) were analyzed with mixed-effects logistic regression which uses a logit transformation of the binary dependent variable. Again, intercepts for these models were allowed to vary by the subjects and words as random factors. The slope for trial number was allowed to vary by subject to account for fatigue or practice effects. P-values for this set of models were computed using the Wald Z statistic.

Skipping statistics were not analyzed due to the dearth of skipped target words (by design).

4.3 RESULTS

All estimated model coefficients and p-values are summarized in Table 7. Results are discussed in terms of patterns present in the fitted model.

4.3.1 Main Effects of Training

There were no main effects of orthographic or meaning training in any of the measures, though there was a trend toward increased first fixation durations with orthographic training ($p < .06$). The addition of a phonology training component resulted in a higher probability of refixating the word on the first pass and (in part because of that effect) longer gaze durations. There was also a main effect of the number of training exposures on the probability of refixating, the probability of re-reading, gaze duration, and total viewing time, with more training exposures resulting in lower likelihoods of refixations and re-reading as well as shorter gaze durations and total viewing times. Although there were few main effects of the training conditions, there were several significant interactions between the training conditions and individual difference measures which reveal that the training can have opposite effects depending on the skill and expertise of the learner. There were also interactions between the training conditions and the number of exposures, with some effects of training only occurring with several exposures. The effects of the each training condition, including these interactions, are described below.

Table 7: Study 3 Model Coefficient Estimates

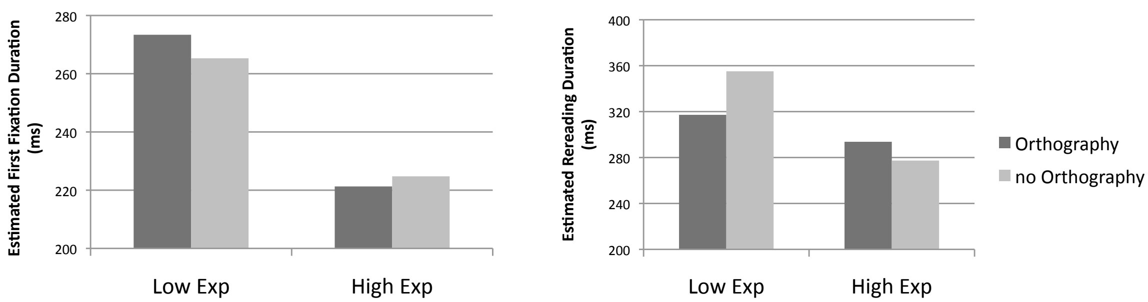
					Initially Viewed Words Only			
		TVT	P(Refix)	P(Reread)	FFD	Refix	GD	Reread
	Intercept	6.160***	-0.029	-0.883***	5.505***	5.523***	5.916***	5.721***
	Trial Number	-0.002***	-0.003***	-0.008***	0.000	-0.002***	-0.001***	-0.002***
Training Factors	No Orthography	0.027	-0.089	0.214	-0.030	0.014	-0.034	0.113
	No Phonology	-0.026	-0.384*	-0.048	-0.018	0.024	-0.070*	0.061
	No Meaning	0.011	-0.151	-0.079	0.014	0.063	0.003	0.098
	# Exposures	-0.032***	-0.080*	-0.104*	-0.005	-0.022	-0.023**	-0.013
Individual Differences	Expertise	-0.038	-0.034	-0.108	-0.044	-0.002	-0.031	-0.016
	Sublexical	-0.260***	-0.626**	-0.560*	-0.078**	-0.191**	-0.211***	-0.103
	Learning/Memory	-0.145*	-0.392*	-0.419	-0.032	-0.091	-0.112*	-0.010
	Accuracy Focus	0.085	0.036	0.404*	0.039	0.045	0.043	-0.040
Interactions	Expertise x Orthography	-0.009	0.020	0.047	0.019	-0.031	0.008	-0.071
	Expertise x Phonology	0.003	-0.002	0.063	0.024*	-0.019	0.014	-0.048
	Expertise x Meaning	0.029*	-0.035	0.179**	0.009	0.024	0.008	0.002
	Orthography x Sublexical	-0.011	0.052	-0.098	0.017	0.011	0.025	-0.026
	Phonology x Sublexical	-0.023	0.084	-0.190*	0.026*	-0.037	0.032*	-0.016
	Meaning x Sublexical	0.020	0.082	0.027	0.003	0.006	0.017	0.017
	Expertise x # Exposures	-0.003	-0.012	-0.018	-0.004	0.005	-0.004	0.021*
	Sublexical x # Exposures	0.007	0.030	-0.008	-0.004	0.015*	0.007	-0.003
	# Exposures x Learning/Memory	0.002	-0.010	0.012	0.005	-0.002	0.000	-0.013
	Orthography x # Exposures	0.022*	0.077	0.002	0.008	0.037*	0.029**	-0.018
	Phonology x # Exposures	0.011	0.104*	0.020	0.000	0.007	0.020*	-0.016
	Meaning x # Exposures	0.007	0.047	0.044	-0.005	-0.002	0.003	-0.007

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

4.3.2 Orthography

Exposure to the visual form of the word during training surprisingly showed a trend toward increasing first fixation durations ($p < .06$), especially for readers with less expertise ($p < .06$ for the interaction). Readers with more expertise had shorter first fixation durations when orthographic training was added. Orthographic training also trended towards producing decreased durations of rereading times, but only for low expertise readers ($p < .07$ for the interaction). Thus, a pattern emerged whereby orthographic training for low expertise readers increased first fixation durations but decreased rereading durations. (See Figure 7).

Figure 7: Interaction Between Orthography Training and Reader Expertise



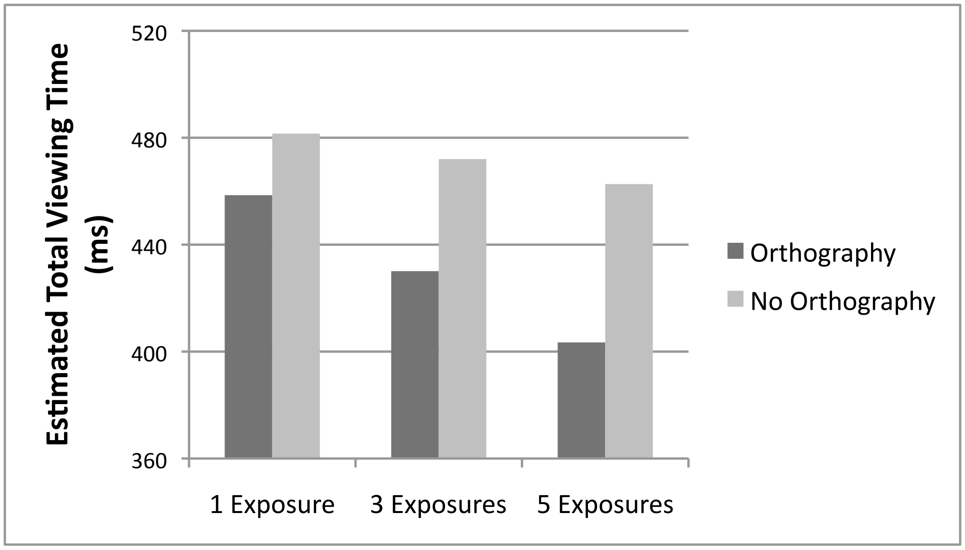
Model predictions for the bounding expertise values illustrating the interaction effect: Less expert readers show longer first fixation durations but shorter rereading durations with orthographic training.

Predictions are back-transformed from log fixation durations for display in msec.

Conditions in which orthography was learned showed a stronger training effect – shorter fixation durations with increasing training exposures – compared to conditions in which orthography was not learned, which did not produce shorter fixations (or showed the opposite pattern) with more exposures. This pattern was evident in the duration of first-pass refixations, the gaze duration, and the total viewing times. The pattern is shown in Figure 8 for total viewing times.

Overall, the patten was that orthographic training decreased the durations of fixations, but did not strongly affect probabilities of refixating or re-reading. Effects were stronger with more training exposures, and were influenced by individual difference scores on the expertise measure. Less-experienced readers showed the trend of having somewhat longer first fixation

Figure 8: Effect of Orthography Training on Total Viewing Time



Model predictions illustrating the interaction between orthographic training and number of training exposures.

Predictions are back-transformed from log total viewing times for display in msec.

durations with the addition of orthographic training, but shorter re-reading durations.

4.3.3 Phonology

The addition of word pronunciation training resulted in an interesting general pattern in which reading was often less efficient (longer fixation durations and more fixations), but generally only when there were fewer training exposures or when the reader skill was low along some dimension. The decrease in efficiency tended to disappear or reverse with more training or for readers of higher skill. This pattern was true across a number of measures.

Pronunciation exposure interacted with reader expertise and sublexical skills in its effect on first fixation durations: more-experienced readers and those with good sublexical skills showed no effect or had shorter durations after pronunciation exposure whereas low expertise readers and those with poor sublexical skills were slowed down by the additional phonology training.

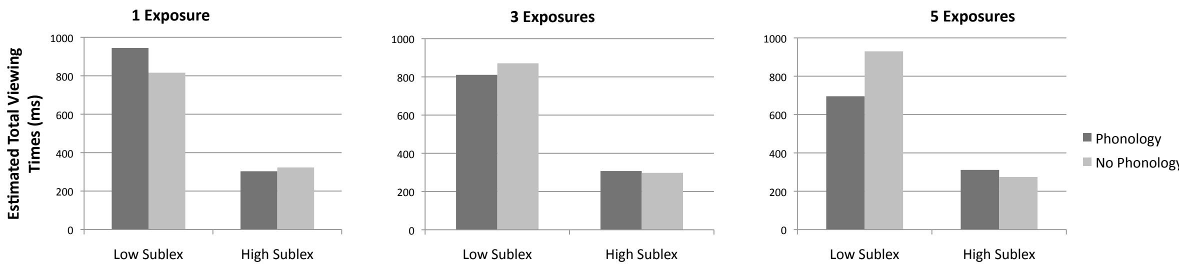
Pronunciation training increased the probability of refixating words on the first pass and increased gaze durations. Although this reached significance as a main effect, an interaction between phonology training and number of exposures showed that the probability of refixating was higher only when phonology was trained with fewer exposures, and was lower when phonology was trained with more exposures. The same interaction was true for gaze duration as well: phonology training increased gaze durations only when phonology was trained with fewer exposures, but was decreased when phonology was trained with more exposures (this may be due to the pattern of refixation probability). In addition, gaze durations were only longer with phonology training for those with lower sublexical skills.

Finally, the addition of phonological training reduced the probability of re-reading (second pass reading), but only for readers with low sublexical skills (those with good sublexical skills already showed a much lower probability of re-reading in all conditions).

Overall, the addition of phonology training seemed to initially slow down reading via longer first fixation durations, increased probabilities of refixating on the first pass, and increased gaze durations, especially when there were fewer training trials and especially for those readers with lower sublexical skills or who were slower, less experienced readers. This extra time spent on the first pass seemed to decrease the probability of re-reading the word later compared to when phonology was not explicitly trained, and did not, therefore, affect total viewing times.

Based on these findings, a three-way interaction between sublexical skill, phonology training, and number of exposures was added to the models. This three-way interaction was significant for probability of refixating, gaze duration, probability of rereading, and total viewing times. The nature of the interaction was consistent with the results reported: readers with high sublexical skills were faster overall and showed little effect of the phonology training. Readers with low sublexical skills showed a slow-down with phonology training after one exposure, but by five exposures they showed a benefit. This three-way interaction is shown for total viewing time in Figure 9, which also demonstrates the general trend of the findings for phonology training.

Figure 9: Effects of Phonology Training and Sublexical Skill on Total Viewing Time

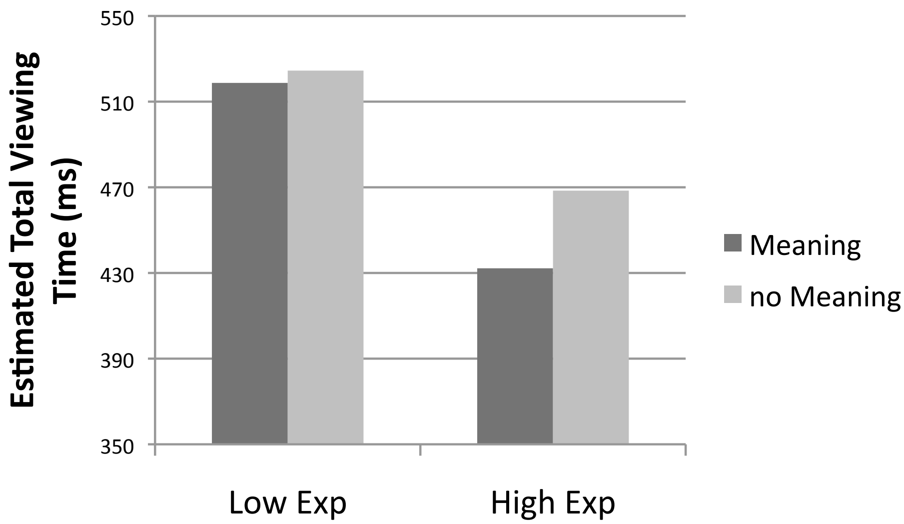


Graph of model predictions for the bounding sublexical skill values illustrating the interaction effect. Predictions are back-transformed from log total viewing times for display in msec.

4.3.4 Meaning

Meaning training significantly affected the probability of re-reading, but did not affect any of the first pass measures of reading. The direction of the effect, similarly to the pattern seen with phonology training, depended on the reader's expertise score. Less-experienced readers were more likely to re-read when meaning was trained. More-experienced readers were less likely to re-read when meaning was trained. This pattern carried over to the total viewing times, but in the case of total viewing times, the pattern was more suggestive of experienced readers having reduced total viewing times with the meaning training, with little effect for less-experienced readers (See Figure 10).

Figure 10: Effects of Meaning Training and Reader Expertise on Total Viewing Time



Model predictions illustrating the interaction between meaning training and reader expertise. Predictions are back-transformed from log total viewing times for display in msec.

4.3.5 Individual Differences

4.3.5.1 Expertise Unlike in Study 2, a reader's score along the Experience/Speed dimension had limited effects on reading behavior for trained words. As reported in the training modality results, the more expert readers were facilitated with orthographic training on first fixation durations, and from meaning training in the probability of re-reading and in total viewing times, moreso than less-experienced readers did (who showed the opposite effect, if anything). One additional effect of a reader's expertise was an interaction with the number of training exposures on the time spent re-reading (in the cases when re-reading occurred). Only the less expert readers showed reductions in re-reading durations with more training.

4.3.5.2 Sublexical Skills Sublexical skills seemed to be the most important individual difference in predicting reading behavior in Study 3. Besides the multiple interactions between sublexical skills and phonology training, those with better sublexical skills had shorter first fixation durations, less probability of refixating, shorter refixation durations, shorter gaze durations, less probability of re-reading, and shorter total viewing times.

4.3.5.3 Learning/Memory Learning/memory interactions with specific training conditions were not included in the model, but those with better learning and memory skills showed lower overall likelihoods of refixating, shorter gaze durations, and shorter total viewing times. This effect likely reflects better learning of or memory for the target words. The effect did not interact with the number of training exposures.

4.3.5.4 Accuracy Focus Participants with lower levels of accuracy in the battery of tests showed a decreased probability of re-reading. These may have been more careless readers who did not re-read when they might have needed to. This pattern may be more reflective of a reading strategy than a skill or knowledge.

4.4 CONCLUSIONS

Results from this experiment highlighted that each component of word knowledge has a distinct effect on reading behavior, and that one way that the components differ is in the temporal locus of their effects. Visual familiarity or orthographic knowledge affected the durations of first pass fixations, as early as the first fixation duration, and sometimes the duration of re-reading fixations. It did not affect the probabilities of re-fixating or re-reading. The auditory training also had an effect on first fixation durations, as well as the probability of refixating and the probability of re-reading, but not the durations of those subsequent fixations. The combination of first fixation effects and first pass refixation probabilities resulted in an effect on gaze duration. The addition of meaning training affected only the re-reading probabilities, but not first pass measures.

This pattern of results suggest that form information is the most important determiner of first pass fixations, with meaning training becoming important in determining re-reading behavior. This supports the idea that readers rely heavily on form information for determining when and where to fixate on the first pass, and that generally these decisions give enough time for meaning processing, because words with more familiar forms tend to also have well-known meanings. When there is difficulty with meaning access, readers must look back and re-read the word.

Additional information given during training, however, did not always increase reading speed. Those with poor sublexical skills were slowed down by the addition of phonology training in their first fixation durations, probability of re-reading, and gaze durations. Less expert readers were also slowed down by the additional pronunciation training in their first fixation durations. In addition, fewer exposures to the pronunciation increased the probability of refixating and the gaze durations compared to more exposures. It may be that the new information, especially with fewer exposures or for those with low sublexical skills, was not yet incorporated into a more context-independent word representation. For example, in a condition in which the phonology was not trained, a reader might quickly produce a rough approximation of the word's pronunciation with no "check" about whether the pronunciation is correct or not. After a single auditory exposure, the reader may recall the prior training episode when they encounter the word in text, or, especially for less skilled decoders, may decode the word in a way that is a mismatch with the trained pronunciation. Accessing specific episodic traces and/or mismatches between generated and learned pronunciations could slow the reader down. After more exposures, especially for better decoders, the word pronunciation may become better incorporated into the lexical representation for that word, and would be more quickly and accurately accessed when the word is encountered. Several exposures to the pronunciation of a word could help establish the phonological representation so that it is accessible from the written word form. Those with poor sublexical skills may need more exposures to a word's pronunciation than those with better sublexical skills before they are able to form a stable phonological representation.

These results are consistent with the framework of the Word Experience Model (Reichle & Perfetti, 2003). In this model, word knowledge is built from the accumulation of individual experiences with each word. With more exposures, parts of the experiences with a word that are repeated (such as the word form and meaning) are reinforced, and parts of the experiences that are specific to a particular exposure (such as the context) do not become part of the abstract word representation. As context-free, abstracted word representations become stronger, information about the word can be accessed more quickly and accurately. The current study suggests that within the word experience framework, more skilled readers may be able to establish context-independent representations with fewer exposures to words than less skilled readers require. This could be achieved through the ability of more skilled readers to establish stronger and more complete episodic traces of each word experience than less-skilled readers, as has been observed via ERP markers of episodic memory following a similar training paradigm (Perfetti et al., 2005; Balass et al., 2010). Strong episodic traces should allow episodic recall of a particular experience to occur quickly and accurately, and it should also allow a more qualitative shift away from this kind of recall of the word learning episode towards context-independent knowledge of the word.

Study 3 results can be summarized as (1) The kind and amount of information about words that readers have been exposed to regulates their eye movements when they read those words, with form information primarily affecting first pass eye movements and meaning information affecting re-reading. (2) Readers' skills determine how their word experience will shape their reading behavior: a "rich get richer" patten is evident in which skilled, experienced readers show more facilitation than less expert readers from the same number of exposures to words. (3) When first learning words, additional components of form training may slow first fixation durations and increase first pass refixations, but ultimately decrease the amount of re-reading that needs to be done. This pattern may indicate that the information is slowly but successfully extracted on the first pass, perhaps through episodic recognition rather than the faster lexical access route. A lack of knowledge may cause readers to initially skim the word, but later re-read the word. This could be a way to avoid disrupting the flow of reading on the first pass.

5.0 STUDY 4: EFFECTS OF PARTIAL WORD KNOWLEDGE ON EYE MOVEMENTS (TESTING PARADIGM)

5.1 INTRODUCTION

Study 3 showed that each component of training (orthographic, phonological, and meaning) had unique effects on eye movements when participants read trained words in text, and that many of the effects varied by the number of training exposures. Although we expected to attribute those effects to differing levels of knowledge and kinds of knowledge about the words, it is possible that some of the training effects, especially decreased reading speed, were a result of participants' episodic memory of the training preceding the reading session. Readers may not have successfully incorporated the new knowledge into their lexical representations. Instead when they see the freshly trained and very rare words in text, they may be recalling, either consciously or subconsciously, the training episodes for that word. More generally, Study 3 may have addressed what happens in the earliest stages of vocabulary acquisition rather than addressing more ubiquitous cases of partial word knowledge or partially deficient word representations.

Like Study 3, Study 4 was designed to examine the direct effects of the knowledge a reader has about a specific word on their eye movements when reading that word. Rather than controlling the amount and kind of knowledge a reader has about a word through an experimental training paradigm, words that were likely to have degraded representations along at least one dimension (of spelling, pronunciation, or meaning) were chosen via a norming study. Participants read these words incidentally in paragraphs with no knowledge that they were target words as part of Study 2. Participants were brought back later to be tested on their knowledge of the words. It was not assumed that they gained significant knowledge of the words during the paragraph reading session or during the time between that session and the testing session. Participants were not alerted to the fact that they had previously read the target words in the eye-tracking session, and instead thought they were just returning for a separate vocabulary test.

Performance on this subsequent spelling, pronunciation, and definition test was used in a model to predict the reader's eye movements when reading those particular words in context. In this way, we hoped to characterize the direct effect of word knowledge on reading behavior with words that were naturally partially known, rather than trained in the laboratory.

5.2 METHODS

5.2.0.5 Target Words Target words were selected such that participants would be likely to have variable knowledge about the spellings, pronunciations, and meanings. To select such words, a norming study was conducted using a non-overlapping group of participants from a similar population (the University of Pittsburgh community). The purpose of the norming was to test participants' knowledge of a set of 7-8 letter, 3 syllable words.

From these results, words were chosen according to how well-known they were across the dimensions of spelling, pronunciation, and meaning. Words were chosen that were well-known on all dimensions, not-well-known on any dimension, or were well-known on one but not another dimension. For example, words for which the spelling but not the meaning was typically known were chosen. More detail about target word selection is described below.

Possible target words were generated using the MRC database with the following constraints: 7-8 letters, 3 syllables, Kucera-Francis written frequency minimum of 2, exclude capitalization (eliminating proper nouns), include only nouns (comprehensive syntactic category), include only standard words (not rare, archaic, etc.), exclude different stress patterns and different phonology (to reduce pronunciation variability). Foil definitions for a multiple-choice test of the word meanings, were definitions of words that would appear in similar contexts. The possible target words were fed into LSA to produce the 20 nearest neighbors using the "General Reading up to 1st year of College" database with the maximum (300) factors. We required a minimum corpus frequency of 5, and the input text was "term" so that no weighting was used. The original term, along with the 20 nearest neighbors were then fed into WordNet. The noun definitions of the original term and its 20 nearest neighbors were extracted as possible foil definitions. Three of these were selected by hand as foils and modified as necessary for clarity and brevity. If the term itself was included in any definition, the definition was modified so as not to include the term. This resulted in a multiple choice test containing foil definitions that would appropriately define words that could occur in similar contexts. It also ensured that all definitions were dictionary definitions from the same dictionary (WordNet). Words were removed when they did not actually fit the intended criteria (e.g. they were not 3-syllable words, they were not nouns), when no reasonable foils were generated, or if there was not a clear primary meaning (i.e. "cabinet"). The original 655 words were reduced to 498 words through this process.

A test was created containing three sequential questions per word for all 498 words. The first was a four-choice spelling question. Three spelling foils were created by hand. The second was a four-choice meaning question (which did not reveal the correct spelling). Three foils were created as described above. The final question asked participants to rate how confident they were that they could correctly pronounce the word. They were given 4 choices (very confident, somewhat confident, somewhat not confident, or not confident). The overall test was divided into four tests of 100 words each and one test of 98 words.

One-hundred and four native English speakers participated in the norming and were compensated with course credit for the Introduction to Psychology course. Participants spent one hour completing one of the five tests, resulting in approximately 21 participants being tested on each word. Participants were instructed that for the multiple-choice spelling and meaning questions, they were allowed to, and *should* answer with multiple answers if they were uncertain. This way we could distinguish between correct guesses and known answers. We could also determine the degree of uncertainty by accounting for the number of answers marked. If all four answers were marked, the participant was completely uncertain. If a single answer was marked, the participant was completely certain.

From this norming, 101 words were chosen which had the highest and lowest ratios of spelling/meaning, spelling/pronunciation, or pronunciation/meaning confidence, resulting in a set of words for which people had better knowledge along one dimension than another. In addition, words with the highest and lowest scores across all 3 dimensions were chosen. Paragraphs which incidentally contained the target words one time (not as the paragraph topic) were found online. Some paragraphs contained multiple target words, resulting in 68 paragraphs containing the 101 target words.

5.2.1 Procedure

Paragraphs containing the target words were the same paragraphs used in Study 2. A subset of 37 native English speakers from Study 2 was recruited for Study 4, with 18 of them having complete and useable data for both portions. These participants returned a minimum of 11 days after participating in Study 2 to be tested on their knowledge of the target words contained in the paragraphs read in Study 2. They were not told that the Study 2 paragraphs had contained these target words, and no indication was given that the vocabulary test was related to the paragraphs. It was assumed that no significant knowledge of the words was obtained between the Study 2 paragraph reading session and the Study 4 testing session, and that enough time had passed between the two sessions that there would be no explicit memory of having read the words in the paragraphs. This would eliminate any of the episodic effects we may have seen in Study 3 results.

The vocabulary testing was done on computer using E-Prime 2.0. Participants were first tested on spelling knowledge for the full set of words. The same 4-choice spelling test was used as that used in the norming study. Participants were instructed to answer with multiple answers if they had any uncertainty as to the correct spelling. They were told to never guess and to instead be sure that the correct answer was always among their choices. Thus, if they were completely uncertain of how to correctly spell the word, they should choose all four possibilities. If they could narrow it down to two of the possible spellings, they should answer with both choices. They were told to only answer with one choice when they were completely certain that the choice was correct. Five practice trials were given in which the experimenter demonstrated answering with multiple answers.

After the spelling block, participants had a 4-choice definition test, again using the same choices generated for the norming study. Participants were given the same instructions as they were for the spelling block, and were encouraged to answer with multiple answers if they had any uncertainty. The same five practice words were used to demonstrate the definition test portion. The final block was a pronunciation/naming block, with the same five practice words. Participants were instructed to pronounce the words as quickly as possible when they appeared on the screen. Reaction time was measured. Then, they were asked to rate their confidence that they correctly pronounced the words on a scale from 1 to 7, with 7 being "very confident" and 1 being "not confident".

5.2.2 Dependent Variables

Dependent variables were computed in the same manner as in Study 2, and included first fixation duration, refixation duration, gaze duration, re-reading duration, total viewing time, probability of skipping, probability of refixating on the first pass, and probability of re-reading.

5.2.3 Independent Variables

The independent variables included the number of answers chosen in the spelling test, the number of answers chosen in the definition test, whether the correct answer was among those chosen in the spelling and definition tests, the confidence rating for pronunciation, and the log reaction time for naming. Each of these variables was centered by subtracting the mean. In addition, interactions between the number of choices and accuracy were included. Subject and item were included as random factors.

5.2.4 Data Analysis

Data for reading times were analyzed using Linear Mixed Effects Regression (LMER) analyses (Baayen, 2008; Baayen et al., 2008) . Subjects and words were included as random factors. For all models, only the intercept was allowed to vary by the random factors. The models were fit using restricted maximum likelihood. P-values were obtained based on highest posterior density confidence intervals computed using Markov chain Monte Carlo (MCMC) sampling with 10,000 iterations (see Baayen, 2008, p.270). This avoids anti-conservative p-values that can arise from use of the t-statistic with the upper bound of degrees of freedom.

Probabilities of skipping, refixating, and re-reading (binary coded variables) were analyzed with mixed-effects logistic regression which uses a logit transformation of the binary dependent variable. Again, intercepts for these models were allowed to vary by the subjects and words as random factors. P-values for this set of models were computed using the Wald Z statistic.

5.3 RESULTS

5.3.1 Behavioral Results

If the correct answer was included in any of the choices, the item was scored as "correct." The average number of spelling options selected was 1.22 ($SD = .62$), and the average proportion correct was .94($SD = .23$). 85.6% of the items were answered with one answer, indicating maximum confidence. The average number of meaning options selected was 1.33 ($SD = .82$) with a mean accuracy of .93 ($SD = .25$). 82.9% of the words were answered with one answer. The average naming speed was 819.41 ms ($SD = 286.75$), although it should be noted that there were times when the microphone did not register the participant's voice right away. The average pronunciation confidence score was 6.66 ($SD = .93$), with 82.3% of the words answered with the top confidence rating of seven.

The correlation between the number of spelling options selected and the number of meaning options selected was significant at the $p < .001$ level, $r = .407$.

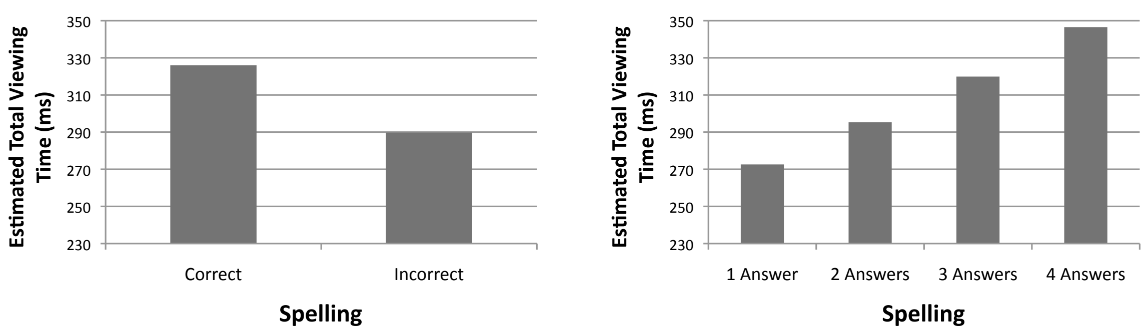
Although words were chosen that were difficult for participants in the norming study along at least one dimension, performances on the spelling, meaning, and pronunciation tests in Study 4 were very high, and a relatively small portion of the words were answered with a lack of confidence. In addition, the confidence in a word's spelling and meaning were significantly correlated, making it difficult to parse effects due to spelling vs. meaning.

5.3.2 Eye-tracking Results

All estimated model coefficients and p-values are summarized in Table 5.3.2. Results are discussed in terms of patterns present in the fitted models.

Less confidence on the spelling test (more options chosen) was associated with longer total viewing times, at least partly a result of a significantly higher probability of re-reading the word. Spelling accuracy, however, showed a surprising pattern in which the accurate words (words with the correct answer included among the selected options) had longer total viewing times, with a higher probability of being re-read. See Figure 11. One possibility is that re-reading the words in the eye-tracking portion may have helped established stronger orthographic representations. It can not be assumed that the paragraph reading session was not a learning session, despite our attempt to delay testing and de-emphasize the target words.

Figure 11: Effects of Orthography Knowledge on Total Viewing TImes



Model predictions for the effects of orthography knowledge on total viewing times. A greater number of answers chosen is associated with less confidence.

Predictions are back-transformed from log fixation durations for display in msec.

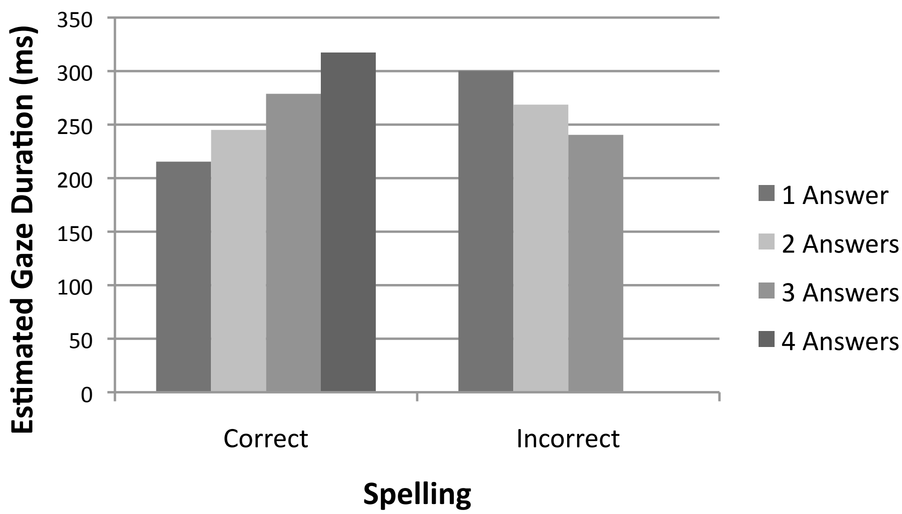
Another possibility is that the incorrect words are the least visually familiar, especially as the number of answers chosen increases. When a reader, for example, chooses three answers, but is incorrect, it means that the only eliminated choice was the correct choice. This means that the correctly spelled word must look "wrong" to the reader, despite the reader's lack of confidence in knowing the correct spelling. These most unfamiliar words might behave

Table 8: Study 4 Model Coefficient Estimates

	Initially Viewed Words Only							
	TVT	P(Refix)	P(Reread)	FFD	Refix	GD	Reread	P(Skip)
Intercept	5.728***	-1.422***	-1.855***	5.350***	5.495***	5.552***	5.670***	-1.805***
Spelling # Answers	0.080**	0.009	0.342*	0.012	-0.021	0.009	0.054	0.077
Spelling Accuracy	0.118*	0.242	0.730*	-0.009	0.181	0.028	0.155	0.304
Meaning # Answers	0.043*	0.007	0.131	0.013	-0.004	0.014	0.095.	-0.030
Meaning Accuracy	0.047	-0.226	0.068	0.052	0.123	0.027	0.127	0.476
Pronunciation Confidence	-0.009	-0.015	0.159	0.008	-0.113**	-0.011	-0.069	-0.019
Naming logRT	0.061	0.326	0.269	0.052	-0.038	0.057	0.047	-0.205
Spelling # Ans x Acc	0.040	0.496	-0.806	0.188.	0.574	0.240.	-0.233	0.441
Meaning # Ans x Acc	-0.041	0.736	-0.798	0.008	-0.049	0.089	-0.068	-0.297

*$p < .05$; **$p < .01$; ***$p < .001$

Figure 12: Interaction Between Accuracy and Confidence of Orthographic Knowledge



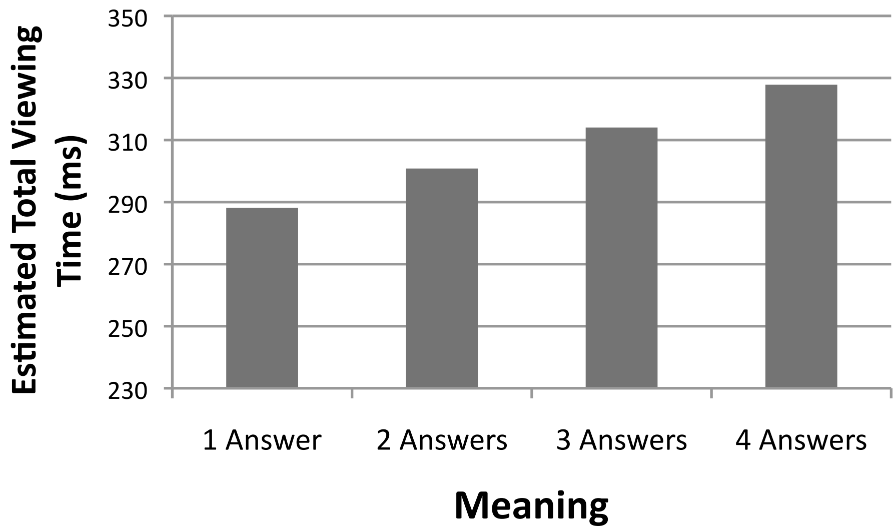
Predicted gaze durations illustrating the interaction between accuracy and confidence of orthographic knowledge. More answers corresponds to decreased confidence.

Predictions are back-transformed from log gaze durations for display in msec.

similarly to the untrained orthography words in study 3: shorter first fixation durations with increased rereading durations. In fact, we find that there was a trend toward an interaction between accuracy and confidence for spelling in first fixation durations ($p < .08$) and gaze durations ($p < .06$). The pattern was that for correct words, a decrease in confidence was associated with longer viewing times. See Figure 12. For incorrect words, in contrast, a decrease in confidence was associated with *shorter* viewing times, perhaps as the words become decreasingly visually familiar.

Less confidence on the definition test (more options selected) was also associated with longer total viewing times (See Figure 13); specifically, there was a trend toward more time spent re-reading words for which the meaning was uncertain ($p < .07$). There were no reliable effects of meaning accuracy and no interactions with meaning accuracy.

Figure 13: Effect of Meaning Knowledge on Total Viewing TIme

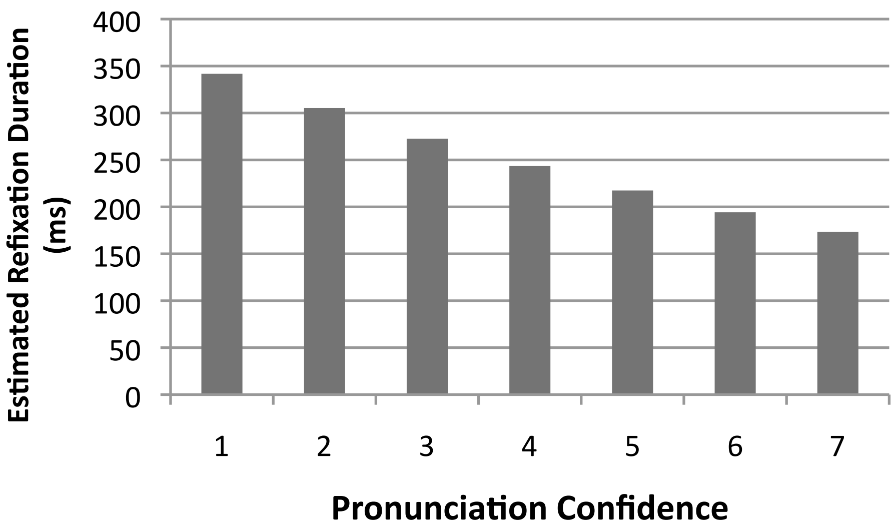


Predicted effect of meaning knowledge confidence on total viewing times. More answers corresponds to decreased confidence.

Predictions are back-transformed from log total viewing time for display in msec.

Higher levels of pronunciation confidence resulted in shorter first pass refixation durations, though this did not translate to reliably shorter total viewing times. (See Figure 14). No effects of naming time were found.

Figure 14: Effect of Pronunciation Confidence on Total Viewing TIme



Predicted effect of pronunciation confidence on refixation durations.

Predictions are back-transformed from log refixation durations for display in msec.

5.4 CONCLUSIONS

Reading measures for the same set of words appearing in the same contexts reflected the individual reader's level of knowledge about the word, even after distributions for the items and readers were taken into account. Both spelling and meaning knowledge affected total viewing times, with less well-known items along each dimension having longer total viewing times.

Orthographic confidence and accuracy affected first fixation durations, gaze durations, the probability of re-reading, and total viewing times. The effects on first pass measures (first fixation durations and gaze durations) showed up as an interaction between confidence and accuracy. Correct words showed increasing first fixation durations (slower reading) with decreasing confidence. Incorrect words showed decreasing first fixation durations (faster reading) with decreasing confidence. This is perhaps because words that are both incorrect and uncertain are arguably the least visually familiar. All of the less familiar words (those for which more answers were chosen) resulted in higher probabilities of re-reading, and longer total viewing times. This result is very similar to the pattern seen for the words in Study 3 that were not trained visually – less time was spent on the first pass, but more time was spent re-reading. In this experiment, though, only the very least visually known words were read more quickly on the first pass. Words with correct-but-uncertain spellings were read like the more well known words on the first pass, but they were later re-read. These words may have been visually familiar enough for the reader to spend an appropriate amount of time reading on the first pass, but unfamiliar enough that they ultimately required re-reading. The re-reading process may actually help the word become more visually familiar in the future – it may be the first step in acquiring a new word representation.

The combination of these findings with the Study 3 findings suggest that when there is insufficient knowledge and a word simply looks unfamiliar, a reader may actually skim a word more quickly than when they have a little bit more familiarity with the word. This strategy may not be the best way to learn the new words in context, but it may be the best way to maintain comprehension when encountering unknown words. Once a word reaches a certain level of familiarity, more time may be spent on the word during reading, and this may be when learning occurs and the word begins to form an accessible representation. Once that representation becomes more solid and stable, words can be read more quickly, but now with access to the stored representations.

Similarly to Study 3, words with meanings that were not well known produced longer total viewing times, at least partly as a result of in increase in re-reading duration. This confirms the later effects of meaning knowledge found in the training study.

Higher levels of pronunciation confidence were associated with shorter first-pass refixation durations. Thus, as in Study 3, first pass fixations were affected by phonological information. However, the durations of first-pass refixations were not specifically affected in Study 3 (rather, the probability of refixating was affected). Weaker phonological representations were associated with longer fixation durations, which is similar to the pattern seen for skilled readers and multiple exposures in the training study. This suggests that indeed, the Study 3 slowing effects seen for words trained with few exposures or for less-skilled readers may be reflecting the fact that the words were just trained, and that the pronunciations may not have been fully incorporated, even weakly, into the word representations.

Overall, Study 4 confirms the findings of the training study. Individual elements of word knowledge have different effects on a reader's behavior when reading those particular words. Generally, first pass fixations are affected by form knowledge and familiarity, and later measures are influenced by meaning knowledge (with lingering effects of orthographic knowledge). In addition, this experiment gives some confirmation to an interesting pattern seen in Study 3: words with the lowest levels of familiarity actually have faster reading times than words with low-to-moderate levels of familiarity. This pattern is seen in the interactions between spelling confidence and accuracy whereby the least well-known words – uncertain and incorrect words – are fixated for less time than other words. Words with higher levels of familiarity are general faster to read than words with low-to-moderate levels of familiarity. This is seen in the overall speed-up in reading measures for words with known spellings, meanings, and pronunciations.

6.0 SUMMARY AND CONCLUSIONS

6.1 SUMMARY OF RESULTS

6.1.1 Study 1: Identifying Individual Differences

Study 1 identified dimensions of variability among adult readers based on a factor analysis of a database of scores on a variety of reading assessments. Five orthogonal dimensions of variability were extracted from the analysis representing (1) reading expertise, (2) sublexical skills, (3) accuracy focus, (4) learning/memory, and (5) "casual reading". The first factor accounted for the most variability and included assessments relating to reading speed, text exposure (especially books), and attitude about reading. These measures are likely to co-vary because of reciprocally causal relationships: readers with good attitudes about reading do more reading and become faster readers, and this in turn makes the process of reading more enjoyable, resulting in good attitudes about reading. The sublexical skill factor included orthographic, decoding, and phonological skills, suggesting that in normal adult readers, orthographic and phonological representations are tightly linked via decoding principles. The accuracy focus factor indicated whether readers prioritized accuracy, even if it was at the expense of speed. The learning and memory factor and the casual reading factor were based on subsets of questions from the Adult Reading History Questionnaire. The learning/memory factor reflected whether the reader had difficulty in school or difficulty remembering instructions, names, dates, etc.. The casual reading factor reflected the amount of internet, newspaper, and magazine reading the reader reported.

6.1.2 Study 2: Relationship Between Individual Differences and Eye Movements

Study 2 examined the effects of individual differences on eye movement behavior during normal paragraph reading. Results showed that individual differences, especially reading expterise, sublexical skills, and accuracy, play a significant role in reading behavior above and beyond properties of the text.

6.1.2.1 Reading Expertise Reading expertise was computed from the factor analysis results with high weightings from the percent of items attempted within the time limits on the Nelson-Denny vocabulary and comprehension tests, the Adult Reading History Questionnaire questions pertaining to reading speed, attitude about reading, and number of books read, and scores on the Author Recognition Test. Participants with high scores along this dimension were more likely to skip words without re-reading them and less likely to refixate or re-read words, and when they did refixate words, they had shorter refixation durations. Those who scored highly on this dimension but had low accuracy on the accuracy dimension tended to read even faster, with fewer refixations and lower probabilities of regressing back to words. It seems likely that this is a pattern of shallow reading where refixations or regressions may be needed for comprehension but are not made.

Readers who scored highly on the expertise dimension showed greater advantages for words that are traditionally more difficult: low-frequency words and words with high bigram frequencies. In addition, these readers showed more facilitation from a word's high frequency neighbors. This pattern suggests that the mechanism by which reading experience leads to faster reading is by increasing the quality of lexical representations across a broader range of words and by increasing knowledge of and connections to other words in the lexicon.

6.1.2.2 Sublexical Skills Sublexical skills were computed from the factor analysis results with high weightings from the spelling recognition test, the phonological awareness test, and a test requiring participants to determine whether nonword letter strings were pronounced like real words. Participants with high sublexical skill scores were more likely to skip words and less likely to refixate words on the first pass. In addition, those with good sublexical skills had shorter first fixation durations for more frequent words, and when more frequent words were skipped, they were less likely to go back and re-read them, signaling that processing was likely complete for the skipped words. This pattern of increased skips and decreased first fixation durations indicates that sublexical skills expedite the early stages of word identification. The fact that this was especially true for frequent words rather than for the infrequent words which might be expected to require more decoding suggests that the pattern is not a result of on-the-fly decoding, but is instead based on an increased ability to quickly recognize the form of frequently experienced words. This is likely to have resulted from the ability of those with good orthographic and phonological skills to unitize words they experience frequently to a greater degree.

6.1.2.3 Accuracy Accuracy scores were computed from the factor analysis results with high weightings from the percent accuracy on the Nelson-Denny vocabulary and comprehension tests and the Raven's Matrices test, and negative weightings from the percent of items attempted on the Raven's Matrices test (indicating a speed-accuracy tradeoff and reflecting the fact that the test questions increase in difficulty). Because accuracy scores are separate from speed scores and come from timed tests, readers with high scores on this factor can be thought of as readers who are capable of answering questions accurately and move through the test at a pace that allows them to do so, whether this is quickly or slowly. Less accurate readers who were also fast and experienced tended to do less refixating and re-reading, perhaps because they were more motivated to get through the task quickly than to read for meaning. Less accurate readers who were also slower and less experienced were more likely to refixate and re-read, perhaps because they were struggling to comprehend the text without the necessary practice to do so efficiently.

6.1.3 Study 3: Word Knowledge (Laboratory Training)

Study 3 was designed to look more closely at the effect of word knowledge on patterns of reading. This was achieved by controlling the type and amount of exposure to new, rare words, and by accounting for individual differences that might affect how well the words were learned. This served as a way to dissociate more clearly the effects of knowledge of particular words vs. the effects of skills which are more generally applicable. In addition, word knowledge was explored by using the Lexical Quality Hypothesis as a more comprehensive framework. Orthographic knowledge, phonological knowledge, and meaning knowledge were explored separately by controlling the visual, auditory, and definition exposure during training, respectively. The quality of knowledge was varied by adjusting the number of exposures during training.

6.1.3.1 Orthographic Knowledge Visual training reduced fixation durations resulting in less time spent refixating, shorter gaze durations, and shorter total viewing times, but this effect emerged only as the number of training exposures increased. There were also trends toward interactions between orthographic training and reader expertise. More experienced readers tended to have shorter first fixation durations with orthography training, indicating that they were able to process the written form of the word more quickly because of the training. Less experienced readers tended to have longer first fixation durations but shorter re-reading times with the addition of visual training. This pattern indicates that visually unfamiliar words were initially read quickly, or perhaps skimmed, by these readers, but later re-read. When visual training had taken place, readers spent more time processing the word initially, and did not have to go back and re-read the word. If we assume that both more exposures and more reading experience result in better orthographic knowledge, these results suggest that there may be a progression in reading patterns based on this knowledge: Very unfamiliar words are fixated only briefly on the first pass, but then later re-read. Familiar but less well-known words are fixated for longer on the first pass, with a lower probability of re-reading. Better known words are fixated briefly on the first pass with lower probability of re-reading. The pattern of initially reading unfamiliar words quickly may enable readers to maintain the greater text meaning in memory without disruption from the unfamiliar word. Re-reading behavior may serve as both a means to double-check whether the word meaning can be extracted and a means of attending to unknown words so that they may be learned.

6.1.3.2 Phonological Knowledge Similarly to the addition of orthography training, the addition of auditory training increased first fixation durations for less experienced readers. The same was true for those with poor sublexical skills: they showed increased first fixation durations and gaze durations with the auditory training, but also decreased probabilities of re-reading. Auditory training also increased the probability of refixating and the gaze duration when there were few exposures (but showed the opposite pattern with more exposures). Again, the pattern seems to indicate that the least phonologically familiar words were re-read, the ones that were somewhat more well-known were read more slowly on the first pass (but not re-read more often), and the most well-known were read more quickly. In this case, training was more effective for both more experienced readers and those with better sublexical skills.

6.1.3.3 Meaning Knowledge Meaning knowledge did not affect first pass measures of reading, but did affect re-reading. Less-experienced readers were more likely to re-read with the addition of meaning training, with little effect on the total viewing time. More-experienced readers were less likely to re-read with the addition of meaning training, and had shorter total viewing times. These results indicated that a lack of meaning knowledge generally manifests in later measures (re-reading), and that more-experienced readers benefitted more from the meaning training than less-experienced readers. Less-experienced readers showed a weaker but similar pattern to that found with visual and auditory training in which they may have taken slightly longer to read the words with additional definition training.

6.1.3.4 Overall Summary Results from this experiment showed that form knowledge affected first pass measures whereas meaning knowledge affected re-reading times. Faster readers with more experience reading were better able to learn from every component of training. In addition, those with good sublexical skills benefitted more from explicit auditory training than those with poor sublexical skills. A general pattern emerged whereby additional but lower quality form knowledge, either as a result of fewer training exposures or less ability to learn quickly from those exposures changed the pattern of reading without necessarily changing total viewing times; the least known words were read more quickly on the first pass but then re-read later, whereas the form-trained but still not well-known words were read more slowly on the first pass, but with less likelihood of re-reading later. The words with best known forms were read more quickly on the first pass. Meaning effects were seen only in the re-reading times, and mainly for more experienced readers. The increases in first pass reading times may have been a result of explicit recall of the training for those words, an issue addressed in Study 4.

Results also showed that participants who had low accuracy scores from the factor analysis did less re-reading, replicating Study 2 results and indicating that these readers may have been not re-reading despite needing to. Unlike Study 2, those who scored highly on the learning and memory factor exhibited a more efficient reading pattern , with lower likelihood of refixating, shorter gaze durations, and shorter total viewing times. Effects of this factor likely emerged in Study 3 as a result of it being a training paradigm. Also unlike Study 2, experience reading had no main effects on eye movement measures outside of differences in learning from the training conditions. This strongly suggests that the benefit from reading experience comes from knowledge of the experienced words rather than only a general speeding up via motor practice or more generalizable text processing skills. However, the ability to learn from exposures to words does seem to be a generalizable skill that improves with more experience and/or a better attitude about reading. Sublexical skills were also important for both learning (in the phonology training portion) and overall reading speed for the new words. It is interesting to note that sublexical skills in Study 2 mainly speeded the early fixation durations for frequent words rather than infrequent words. This finding contrasts with the importance of sublexical skill in determining reading behavior in Study 3, if the trained words are thought of as infrequent words. Perhaps sublexical skills are most important in initial encounters with unfamiliar words, but less important once a word-specific representation has begun to be established.

6.1.4 Study 4: Word Knowledge (Lifelong Learning)

Study 4 examined the relationship between fixation patterns of words in text and the reader's knowledge of those words. Word knowledge was assessed along the dimensions of spelling (meant to test the quality of orthographic representations), pronunciation (testing the quality of phonological representations), and definitions (testing the quality of semantic representations), using the same framework for word knowledge as the training study (Study 3).

Results showed that when a reader is uncertain about a word's orthography, they are more likely to re-read the word, resulting in longer total viewing times. In addition, the least visually known words (words for which the reader was uncertain about the spelling *and* eliminated the correct spelling in a multiple choice test) were actually read more quickly on the first pass. This replicates the Study 3 finding that the least orthographically well-known words were read quickly on the first pass and refixated later. Words with less known pronunciations caused longer first pass refixation durations in comparison to confidently pronounced words. Lastly, words with less known meanings resulted in longer re-reading durations and total viewing times, replicating the general late effects of meaning knowledge seen in Study 3.

6.2 CONCLUSIONS

Experience with reading is important to developing good reading skills. How much reading (especially book reading) a person does is tightly coupled with their reading speed and attitude about reading. It is not a stretch to assume that there is a causal relationship between practice reading and reading speed (though the reciprocal is also true when readers don't compensate for slow reading with increased time spent reading). Reading experience can cause increases in the efficient and accurate processing of words via a variety of mechanisms including increasing the quality of lexical representations readers have for the words they have encountered and improving more generalizable skills that can be applied to new words.

It is not only the amount of experience readers have that causes these improvements in representations and skills, but also the quality of the experience, the ability of the reader to benefit from the experience given the reader's current knowledge and skill set, and the reader's motivation and attention. Although many experiments have examined the impact of various manipulations of the quality of experience to discover what the most supportive kinds of practice are for word learning, fewer studies have examined the interaction between the learning experience and individual traits and skills of the reader.

The studies executed in this thesis create a picture of how lexical processing proceeds based on the quality of knowledge a reader has about a word. If words with very unfamiliar forms are encountered, readers tend to skim them on the first pass quickly, but then re-read the word later. They may initially skim the word so as not to disrupt the meaning flow of the passage by spending too much time on one word. They will then later re-read the word to spend more time determining whether they can access the meaning, or for very unfamiliar words, they may later re-read the word in order to boost visual familiarity with the word for future encounters. If the meaning of the word is not known, re-reading the word after moving on in the text can help establish the word's meaning based on inferences from the surrounding context. This may be the first stage in acquiring word representations. Williams and Morris (2004) found similar results when they observed adult readers learning words from context. Unfamiliar words that were later correct on a forced-choice meaning test were read more quickly on the first pass, but more slowly on the second pass than words that were later incorrect on the test. The authors attribute the increased re-reading duration to the process of connecting the unfamiliar word with the meaning inferred from the informative text following the word. They suggest that the words were read more quickly on the first pass due to a better metacognitive sense of how well the word was known – words that were correctly sensed to be unfamiliar were read more quickly on the first pass. Results from the experiments presented here corroborate their hypotheses.

Once a word is partially acquired, but is still unfamiliar, more time is spent processing the word on the first pass. This reflects the weaker representation quality – more fixations and longer fixations are needed to get to the meaning, but the reader is familiar enough with the word to know that the information will eventually be accessed. Which particular fixations are increased depends on which parts of the word representations are weak – orthographic and phonological information affect first pass fixations (although the particular fixation times vs. fixation probabilities varied across experiments), and phonological and meaning information (most consistently meaning information) affect re-reading times, (again, the specific durations or probabilities were inconsistent). This finding corroborates priming experiments describing the time-course of activation of different components of knowledge during reading, and extends those findings by verifying that the degree of knowledge of each of these components (as determined through both training and evaluations) also affects reading on a similar time-course. Namely, both orthographic and phonological information are accessed quickly, with semantic information retrieved as lexical access occurs (Lee et al., 1999; Perfetti et al., 1988; Perfetti & Bell, 1991; Rayner et al., 1995).

Many studies have previously found effects of overall word knowledge on the efficiency of lexical access through the use of familiarity norms, which are arguably either good reflections of experiential frequency (which would indirectly predict lexical quality) or take into account aspects of the reader's word knowledge besides frequency of experience, such as perceived meaningfulness of the word (Balota, Pilotti, & Cortese, 2001). In fact, Williams and Morris (2004) found that familiarity ratings predicted eye movements above and beyond frequency effects, and that when familiarity was held constant, there were no differences between words with different frequencies. Interestingly, they also compared two conditions which differed in meaning knowledge (determined via a forced-choice meaning test), but which did not differ in either frequency or familiarity. They found no differences in eye fixations between these two conditions, despite the differences in the meaning test performance, and suspected that semantic characteristics were more important for the forced-choice task than the silent reading task because the words were not in context. In Study 4, we found that it was the confidence a reader had of their meaning knowledge, rather than the accuracy of the knowledge, that had a measurable effect on eye movements. Like Williams and Morris, we did not find differences in eye movement patterns based on the accuracy of our multiple choice meaning task. Readers may use the metacognitive awareness that they are unsure of the word's meaning to prompt them to re-read the word, a mechanism similar to the one proposed by Williams and Morris to explain why visually unfamiliar words were read more quickly on the first pass.

The speed of the process of learning words and incorporating them into the lexicon varies as a function of reader skill (in addition to the number and type of exposures). Previous studies that have examined the effects of individual skill on word learning have generally defined a single, a priori dimension of reader skill (such as comprehension ability or learning speed) and have found that more skilled readers are more effective learners (e.g. J. R. Nelson et al., 2005; Balass et al., 2010; Perfetti et al., 2005). We instead took a data-driven approach to discovering the underlying structure in reading performance, and we found that specific individual attributes within that structure affected specific elements of word learning (orthographic, phonological, and semantic).

Those with greater experience reading are faster to incorporate orthographic, phonological, and meaning knowledge into lexical representations. This could be because some of the trained words had actually been encountered enough before to lay the familiarity groundwork for later acquiring a more solid representation. Alternatively, familiarity could have come from just having exposure to more words in general, because other words have similar spellings, pronunciations, and meanings. Another possibility is that these readers simply have more well-connected networks of word meanings and their relationships to word forms, and fitting new words into that rich network results in the reinforcement, strengthening, and stabilizing of new word representations by the existing network.

Those with good sublexical skills were more quickly able to use additional auditory information, specifically, to improve the quality of lexical representations compared to those with lower sublexical skills. Nelson et al. (2005) found that faster learners depend less on the modality in which a word was learned for later recognition. Relevant to Study 3, fast learners were better than slower learners at recognizing a word that they had heard during training when they later saw the word. This could be the result of an increased ability to orthographically recode the phonological information during learning, or it could be that the readers decode well during subsequent reading and are able to make the connection to the learned phonology. In the latter case, both better decoding ability at the time of reading and the ability to create a phonological representation from an auditory exposure during training would aid in recognition. Although the sublexical skills of the individuals were not tested in their study, it may be the case that their faster learners had better sublexical skills, which would support the observations about about how many exposures were required to learn the words to criterion in addition to supporting the possible explanations for why these fast learners were less modality-dependant. One difference between the Nelson et al. study (2005) and the current study is that the training conditions in the current study both with and without phonology always included the visual component (because there was no completely untrained condition with which to compare the P-only condition). Thus, our good decoders may be especially skilled at linking the pronounced word with the written word to form a more robust lexical representation, allowing the word form to be more quickly identified in reading. Regardless of training condition, those readers with good sublexical skills were more fluent readers of all of the newly learned target words. Sublexical skill seems to result not only in more highly unitized representations of frequent words, as shown in Study 2, but also to allow readers to more efficiently turn a few exposures to word forms into accessible representations.

Reading experience and sublexical skills continue to be important differentiators in adult readers, with consequences on their ability to continue learning vocabulary with fewer exposures. And vocabulary knowledge, in turn, has consequences for fluent reading, with less stable word representations along any dimension resulting in less efficient reading on a word-by-word basis. Effects of the learning process and acquisition of word knowledge continue to be important and measurable in even fairly easy paragraph reading. Those who have done more reading and thus have acquired lexical knowledge more effectively are faster, more efficient readers, especially for less-frequent words. Those with good sublexical skills are also more efficient because they seem to have more integrated form representations for frequent words, which allows them to spend less time on the first fixation. Their skill at the early identification of words helps speed them up on first pass measures generally, and increases the probability that they will skip words. Differences in reader skills seem to be more important during the word learning process, and differences in lexical knowledge, resulting from the combination of skills and exposures, seems to be more important for normal reading.

6.2.1 Implications for Understanding Lexical Access and Eye Movement Control During Reading

Models of eye movement control refine our understanding of lexical access in the context of reading by making quantitative predictions about eye fixations and movements on the basis of characteristics of words, text, or even individual readers. All of the major models of eye movement control include word frequency as a predictor of fixation behavior, including the E-Z Reader model (Reichle et al., 2006), the SWIFT model (Richter, Engbert, & Kliegl, 2006), the Glenmore model (Reilly & Radach, 2006), the SERIF model (McDonald, Carpenter, & Shillcock, 2005), and the SHARE model (Feng, 2006). However, as we have demonstrated especially in Study 2, frequency effects are modulated by individual reader skills and experience. A more complete model of eye movement control will need to incorporate individual variability, especially given that reader variability is more nuanced than simply "faster" or "slower", and in fact modulates the relationship between lexical factors and reading behavior.

There may be at least two mechanisms behind the universally modeled frequency effect: (1) Increased knowledge of and exposure to word forms and meanings, and (2) unitization of the word form, enabling larger form units to be processed. A reader with above-average reading expertise may demonstrate faster reading especially for low frequency words compared to less expert readers via the knowledge mechanism. A reader with above average sublexical skills may demonstrate faster reading especially for high frequency words via the unitization mechanism. Because readers can demonstrate expedited word reading for either high or low frequency words depending on their particular reader characteristics, modeling the eye movement behavior of individual readers based on skill would require there to be more than one component of skill (corresponding at least to expertise and sublexical skill), with each component modulating the frequency effect as described.

This is not to suggest that these must be two totally distinct processes or pathways for learning from experience. For example, connectionist models of lexical access may be able to account for these individual differences much as they are able to account for patterns of dyslexia: those with good sublexical skills may have better (self-generated) form feedback during learning, as suggested by the self-teaching hypothesis (Share, 1995) possibly resulting in hidden units that better reflect word form patterns. Those with more reading expertise may be modeled with more word training (higher experiential frequency). This increased word exposure is likely to be more helpful for low-frequency words as high frequency words approach ceiling for the quality of representations.

Another important concept emerges from Studies 3 and 4. Successful modeling of eye movements for the least orthographically and phonologically familiar words may require either a skewed inverse-U shaped curve for processing time based on form familiarity, or a time limit by which a certain threshold of familiarity must be reached for processing to continue. This type of modification would account for the pattern of decreased first fixation durations for the least form-familiar words.

Considering form familiarity brings us to another point. Form knowledge and meaning knowledge affected different components of the eye movement record, with form knowledge generally affecting first pass measures and meaning knowledge generally affecting later measures. This suggests that models of eye movement control during reading could better describe frequency-based effects by using separate estimates of form and meaning familiarity. To the extent that form and meaning familiarity are correlated, a single measure will be successful. However, accounting for partial word knowledge may require these two separate estimates of familiarity. The E-Z Reader model, for example, has two processing stages which are a function of frequency (and predictability), weighted to different extents. Partial word knowledge could be accounted for in this model by using form familiarity estimates for the early stage (called the "familiarity check"), and meaning familiarity estimates for the later stage. The familiarity check stage could account for the lowest familiarity words via the proposed skewed inverse-U function or the threshold of familiarity that must be reached for processing to continue. In addition, the E-Z Reader model could modulate the effects of these frequency components by reader skill along multiple dimensions.

In summary, frequency effects modeled in current models of eye movement control may be better split into form knowledge and meaning knowledge effects, and estimated by measures of form and meaning familiarity norms. Words with especially unfamiliar forms may need special consideration in models, as behavior when reading these words does not seem to follow a linear pattern of familiarity. Lastly, individual differences in the frequency effects need to be modeled in a more nuanced way than simply increasing or decreasing overall speed or the size of the frequency effect.

APPENDIX A

ADULT READING HISTORY QUESTIONNAIRE

Session Number: _______________ Version AA

Subject ID (Last 4 digits of phone):________________________

Reading History Questionnaire

Please circle the number of the response that most nearly describes your attitude or experience for each of the following questions or statements. If you think your response would be between numbers, place an "**X**" on the line. However, only use an "**X**" to make your response if you are halfway between two numbers. An "**X**" should only be used to indicate a 0.5, 1.5, 2.5 or 3.5.

1. Which of the following most nearly describes *your* attitude toward school when you were a child:

Loved school; Favorite activity — Hated school; Tried to get out of going

0 ________ 1 ________ 2 ________ 3 ________ 4

2. How much difficulty did *you* have *learning* to *read* in elementary school?

None — A great deal

0 ________ 1 ________ 2 ________ 3 ________ 4

3. How much extra help did *you* need when learning to read in elementary school?

Help from: No help	Friends	Teachers/ parents	Tutors or special class *1 year*	Tutors or special class *2 or more years*
0	1	2	3	4

4. Did you ever reverse the order of letters or numbers when you were a child?

No — A great deal

0 ________ 1 ________ 2 ________ 3 ________ 4

5. Did you have difficulty learning letter and/or color names when you were a child?

No — A great deal

0 ________ 1 ________ 2 ________ 3 ________ 4

6. How would you compare *your* reading skill to that of others in your elementary classes?

Above Average — Below Average

0 ________ 1 ________ 2 ________ 3 ________ 4

7. All students struggle from time to time in school. In comparison to others in your classes, how much did you struggle to complete your work?

Not at all	Less than most	About the same	More than most	Much more than most
0	1	2	3	4

8. Did you experience difficulty in high school or college English classes?

No; enjoyed and did well		Some		A great deal; Did poorly
0	1	2	3	4

9. What is your current attitude toward reading?

Very Positive				Very Negative
0	1	2	3	4

10. How much reading do you do for pleasure?

A great deal		Some		None
0	1	2	3	4

11. How would you compare your current reading speed to that of others of the same age and education?

Above average		Average		Below average
0	1	2	3	4

12. How much reading do you do in conjunction with your work? (If you are a full time student, you can consider that your job)

A great deal		Some		None
0	1	2	3	4

13. How much difficulty did you have *learning* to *spell* in elementary school?

None		Some		A great deal
0	1	2	3	4

14. How would you compare *your* current spelling to that of others of the same age and education?

Above average		Average		Below average
0	1	2	3	4

15. Did your parents ever consider having you repeat any grades in school due to academic failure (not illness)?

No	Talked about it, But didn't do it	Repeated 1 grade	Repeated 2 grades	Dropped out
0	1	2	3	4

16. Do you ever have difficulty remembering people's names or names of places?

No				A great deal
0	1	2	3	4

17. Do you have difficulty remembering addresses, phone numbers, or dates?

No				A great deal
0	1	2	3	4

18. Do you have difficulty remembering complex verbal instructions?

No				A great deal
0 ______	1 ______	2 ______	3 ______	4

19. Do you currently reverse the order of letters or numbers when you read or write?

No				A great deal
0 ______	1 ______	2 ______	3 ______	4

20. How many books do you read for pleasure each year?

More than 10	6-10	2-5	1-2	None
0 ______	1 ______	2 ______	3 ______	4

21. How many magazines do you read for pleasure each month?

5 or more	3-4 regularly	1-2 regularly	1-2 irregularly	None
0 ______	1 ______	2 ______	3 ______	4

22. Do you read daily (Monday – Friday) newspapers?

Every day	Once a week	Once in a while	Rarely	Never
0 ______	1 ______	2 ______	3 ______	4

23. Do you read a newspaper on Sunday?

Completely Every Sunday	Scan Each week	Once in a while	Rarely	Never
0 ______	1 ______	2 ______	3 ______	4

24. Do you read newspaper and/or magazine articles on the internet?

Every day	Once a week	Once in a while	Rarely	Never
0 ______	1 ______	2 ______	3 ______	4

*****Check the most appropriate answer for each of the following questions*****

25. To the best of your knowledge, did *your* parents ever report that either one of them had a problem with reading or spelling?

___________ Yes
___________ No
___________ Not Sure

If yes, please give details: ____
____
____
____

26. To the best of your knowledge, did *your* brothers and/or sisters ever have a problem with reading or spelling?

____________ Yes
____________ No
____________ Not Sure

If yes, please give details: ____
____
____
____

27. What is the highest educational level that you have attained?

____________ High school, did not graduate
____________ High school graduate
____________ Trade or business school
____________ Some college, have not graduated
____________ Junior college graduate, associate's degree (or equivalent)
____________ College graduate, bachelor's degree (or equivalent)
____________ Some postgraduate education, no advanced degrees
____________ Attained 1 or more advanced degrees

APPENDIX B

AUTHOR RECOGNITION TEST

Session Number: _______________ Version AA

Subject ID (Last 4 digits of phone):________________________

Author Recognition Test

Below you will see a list of 80 names. Some of the people in the list are popular writers (of books, magazine articles, and/or newspaper columns) and some are not. Please read the names and put a check mark next to the names of the individuals you know to be writers. Do not guess, but only check those who you know to be writers. Remember, some of the names are people who are not popular writers, so guessing can easily be detected.

- [ ] Gustav Mahler
- [ ] T.S. Eliot
- [ ] Stephen King
- [ ] Dan Brown
- [ ] J.K. Rowling
- [ ] Frederick Chopin
- [ ] Thomas Friedman
- [ ] Emilio Pucci
- [ ] Lewis Carroll
- [ ] Eriq LaSalle
- [ ] Joan Miro
- [ ] Tom Wolfe
- [ ] John Grisham
- [ ] William Borah
- [ ] John Singer Sargent
- [ ] Ambrose Bierce
- [ ] Tom Stoppard
- [ ] Mary Higgins Clark
- [ ] Mikhail Baryshnikov
- [ ] Samuel Beckett

- [ ] Sue Grafton
- [ ] George Orwell
- [ ] Angela Lindvall
- [ ] Edward Hopper
- [ ] Anne Rice
- [ ] Andrew Jackson
- [ ] John Constable
- [ ] Mary Shelley
- [ ] Gustav Klimt
- [ ] Helmut Lang
- [ ] Henry James
- [ ] John Irving
- [ ] Isaac Asimov
- [ ] Mary Cassatt
- [ ] Ezra Pound
- [ ] Edgar Rice Burroughs
- [ ] J.M.W. Turner
- [ ] Will Sasso
- [ ] Thomas Mann
- [ ] Eli Whitney

- [ ] V.C. Andrews
- [ ] Laura Dern
- [ ] Leonid Pasternak
- [ ] Henry David Thoreau
- [ ] Danielle Steele
- [ ] James King
- [ ] Annie Oakley
- [ ] Upton Sinclair
- [ ] Martina Hingis
- [ ] David Ortiz
- [ ] Louisa May Alcott
- [ ] Arthur C. Clarke
- [ ] Thomas Eakins
- [ ] George Eliot
- [ ] James Cagney
- [ ] Mike Stanton
- [ ] Primo Levy
- [ ] Charles Spiegel
- [ ] Vivica A. Fox
- [ ] Harriet Tubman

- [ ] Walter E. Kurtz
- [ ] Stephen Crane
- [ ] Ellen Tracy
- [ ] Paul Pierce
- [ ] Chester A. Arthur
- [ ] Tennessee Williams
- [ ] Joseph Conrad
- [ ] Igor Stravinsky
- [ ] Christine Lahti
- [ ] Ansel Adams
- [ ] William Golding
- [ ] Leo Tolstoy
- [ ] Aaron Burr
- [ ] Ernest Hemingway
- [ ] Amy Tan
- [ ] John Updike
- [ ] Nicholas Sparks
- [ ] Linda Carter
- [ ] William Wallace
- [ ] Richard Wright

APPENDIX C

PHONOLOGICAL AWARENESS TEST

Session Number: ________ Version AA
Subject ID (Last 4 digits of phone): __ Date: _____

Phonological Awareness Test

This is a test of your ability to think about how words sound, regardless of how they're spelled. Each question has a word. Say this word silently to yourself. Ex. CRUISE (say to yourself, /krooz/).

Then look at the sound to be removed from the word. Say the word silently to yourself, without the sound. Write the word that you said. Ex. Remove the /z/ sound. (say to yourself, /kroo/. Write the word crew.)

Then look at the sound to be replaced **in the spot where you removed the first sound**. Say the word silently to yourself, adding this sound. Write the word that you said. Ex. Add the /l/ sound. (say to yourself, /krool/. Write the word cruel.)

Here's another example.

SPEAK	remove the /p/ sound	seek	add the /l/ sound	sleek

Sometimes you might actually have to split apart the sounds made by one letter.

MIXED	remove the /k/ sound	missed	add the /d/ sound	midst

Sometimes you might have to put together sounds made by several letters.

PHONICS	remove the /f/ sound	onyx	add the /t/ sound	tonics

*****Hint, Hint: Everything you write should be a real word, correctly spelled.*****

Now complete the following items:

1. MIDDLE	remove the /d/ sound	_______________	add the /s/ sound	_______________
2. QUEEN	remove the /w/ sound	_______________	add the /l/ sound	_______________
3. NICKEL	remove the /k/ sound	_______________	add the /b/ sound	_______________
4. HATCHED	remove the /tch/ sound	_______________	add the /k/ sound	_______________
5. QUAKE	remove the **first** /k/ sound	_______________	add the /uh/ sound	_______________
7. MOTION	remove the /m/ sound	_______________	add the /l/ sound	_______________
8. WRAPPED	remove the /r/ sound	_______________	add the /t/ sound	_______________
9. CAUGHT	remove the /k/ sound	_______________	add the /b/ sound	_______________
10. PAGE	remove the /j/ sound	_______________	add the /n/ sound	_______________
11. LAUGHTER	remove the /l/ sound	_______________	add the /r/ sound	_______________
12. SKY	remove the /k/ sound	_______________	add the /p/ sound	_______________
13. FARCE	remove the /s/ sound	_______________	add the /m/ sound	_______________
14. COLONEL	remove the /n/ sound	_______________	add the /d/ sound	_______________
15. ROD	remove the /r/ sound	_______________	add the /k/ sound	_______________
16. RACKS	remove the /r/ sound	_______________	add the /t/ sound	_______________
17. MIGHT	remove the /t/ sound	_______________	add the /s/ sound	_______________
18. SCORE	remove the /k/ sound	_______________	add the /t/ sound	_______________
19. YACHTER	remove the /y/ sound	_______________	add the /p/ sound	_______________
20. SLIME	remove the /m/ sound	_______________	add the /ss/ sound	_______________

PLEASE DO NOT LEAVE ANY ITEMS BLANK. MAKE YOUR BEST GUESS EVEN IF YOU'RE UNSURE ABOUT THE CORRECT ANSWER.

APPENDIX D

REAL WORD TEST

Session Number: _______________ Version AA

Subject ID (Last 4 digits of phone):________________________

Real Word Test

Below you will see a list of 168 letter strings. Some of the strings in the list, if you sound them out, sound like real words (for example 'gote' sounds like 'goat') – even if they are not spelled like real words. Some of the strings, when you sound them out, do not sound like real words. Please read the strings and check each one that sounds like a real word. Do not guess, but only check those you know to be words. Remember, some of them are made-up words, so guessing can be easily detected.

☐kake	☐lote	☐eckseed	☐daighbyou	☐afe	☐serkyouler
☐plog	☐kalone	☐bloe	☐beal	☐neech	☐muhsheene
☐deace	☐teech	☐spighn	☐bote	☐balit	☐howceholed
☐gairubteigh	☐widome	☐speschelist	☐baionette	☐strale	☐stuckottee
☐warld	☐derty	☐trastor	☐skwees	☐bair	☐ait
☐fite	☐plice	☐horl	☐reash	☐carn	☐thurtone
☐thair	☐jimpozyum	☐craul	☐typoal	☐rakk	☐leeve
☐oquerd	☐muroseize	☐edepust	☐thord	☐broan	☐tracter
☐keitreeot	☐urthkwaik	☐shatcheure	☐jadwarr	☐slotayshin	☐kworrty
☐jope	☐broave	☐phrantcheyes	☐ohoodeoh	☐exkspedeinter	☐feem
☐feefail	☐flote	☐quaynt	☐ferst	☐foap	☐moatev
☐fether	☐strate	☐filce	☐bape	☐mistenk	☐shate
☐mijen	☐shart	☐fearce	☐viknaime	☐gome	☐naim
☐thrue	☐timnastix	☐baik	☐werld	☐peece	☐udderents
☐joak	☐nerr	☐crail	☐traif	☐nowairt	☐nade
☐thurd	☐seaf	☐dofter	☐youroveu	☐dorty	☐fipe
☐neintoon	☐amiture	☐eap	☐indiseeze	☐klass	☐braive
☐dak	☐anteik	☐phoarbarense	☐poantil	☐fead	☐crysieze
☐grafight	☐roak	☐konseed	☐boaf	☐fense	☐filst
☐fither	☐ploor	☐gaim	☐caim	☐meave	☐pame
☐tirn	☐cliss	☐pebl	☐kinzeanyal	☐reech	☐hoap
☐colesture	☐ulsir	☐threp	☐kard	☐docter	☐rondeyvoo
☐skwolb	☐heetir	☐aknoleque	☐lait	☐meer	☐dufay
☐hoate	☐herowique	☐matcheune	☐onsahmbel	☐bodvill	☐anekkdoat
☐kwartet	☐turt	☐fairce	☐floar	☐hawl	☐gudnice
☐eer	☐puteightore	☐saip	☐shaip	☐seashilly	☐shurt
☐theer	☐vazhyoual	☐saif	☐seet	☐firoh	☐futeeg
☐plaice	☐trane	☐skwyde	☐broun	☐urb	☐voastij

APPENDIX E

SPELLING TEST

Session Number: _______________ Version AA

Subject ID (Last 4 digits of phone):________________________

Spelling Test

Below you will see a list of 140 words. Some of the words are spelled correctly, and some are not. Please read the words and check each one that is spelled correctly. Do not guess, but only check those you know to be correctly spelled words. Remember, some of them are made-up spellings, so guessing can be easily detected.

☐tape	☐wimp	☐innitiate	☐chooze	☐streem	☐Caesar
☐schedule	☐inoculate	☐bureaucracy	☐essence	☐obedient	☐lisence
☐take	☐conscious	☐lurn	☐indispensible	☐supersede	☐adviser
☐personnel	☐sircus	☐practicle	☐grevious	☐sheep	☐wize
☐concensus	☐writing	☐official	☐toward	☐persuade	☐minuscule
☐exhileration	☐smoak	☐diction	☐wurd	☐emptey	☐lake
☐collumn	☐vinaigrette	☐cloun	☐expresso	☐nieve	☐occurred
☐guidance	☐hight	☐scair	☐crucifiction	☐buoyant	☐roar
☐equipped	☐decieve	☐certin	☐basically	☐Febuary	☐questionaire
☐embaress	☐businiss	☐priviledge	☐exsaust	☐legible	☐portray
☐dreem	☐wait	☐fue	☐fase	☐inseperable	☐neice
☐astronaut	☐appropriate	☐similar	☐sleap	☐fiery	☐answer
☐nuisence	☐mileage	☐hurt	☐ajoin	☐need	☐room
☐accomodate	☐morgage	☐pastime	☐congradulations	☐please	☐strech
☐wroat	☐forfeit	☐harrass	☐grammer	☐marriage	☐heavy
☐chief	☐amature	☐trew	☐turtle	☐occurrence	☐skaite
☐environment	☐arctic	☐hole	☐mispelled	☐between	☐definite
☐sherbert	☐keap	☐deep	☐goat	☐stoar	☐perseverance
☐agression	☐noticeable	☐absorption	☐mirror	☐foriegn	☐enemy
☐eminent	☐fourty	☐rhythim	☐street	☐recurring	☐governor
☐adultery	☐millennium	☐bowl	☐easy	☐mischevious	
☐pharmacy	☐judgment	☐shoulder	☐young	☐seperate	
☐asterisk	☐thum	☐nise	☐bote	☐rane	
☐beleav	☐wagun	☐snow	☐tragedy	☐evry	

References

Allington, R. (1984). Content coverage and contextual reading in reading groups. *Journal of Literacy Research*, *16*(2), 85–96.

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 802–814.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 234–254.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic bulletin & review*, *4*(4), 439–461.

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General*, *139*(2), 299–318. Available from http://www.sciencedirect.com/science/article/B6X07-504W7FG-5/2/bddcc9f1c9e25b9b7745486e7f0a9e1c

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R.* Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Balass, M., Nelson, J., & Perfetti, C. A. (2010, Jan). Word learning: An erp investigation of word experience effects on recognition .... *Contemporary Educational Psychology*. Available from http://linkinghub.elsevier.com/retrieve/pii/S0361476X1000007X

Balota, D. A., & Chumbley, J. (1984). Are lexical decisions a good measure of lexical access? the role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 340–357. Available from http://www.sciencedirect.com/science/article/B6X08-4NV6PPM-2/2/3ecef6f19ef33e2a38b6c4fd7c566021

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001, Jan). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, *29*(4), 639–647.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The english lexicon project. *Behavior Research Methods*, *39*(3), 445-459 N1 –. Available from http://brm.psychonomic-journals.org/content/39/3/445.abstract

Barry, C., Morrison, C., & Ellis, A. (1199). Naming the snodgrass and vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *The Quarterly Journal of Experimental Psychology A*, *50*, 560–585. Available from http://dx.doi.org/10.1080/027249897392026

Beck, I., Perfetti, C. A., & McKeown, M. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, *74*(4), 506–521. Available from http://www.sciencedirect.com/science/article/B6WYD-4NV6XBD-5/2/cf4fdb9e5bd7978c2015dcb36c1326e9

Biederman, G. (1966). Supplementary report: The recognition of tachistoscopically presented five-letter words as a function of digram frequency. *Journal of Verbal Learning and Verbal Behavior*, *5*(2), 208–209.

Biemiller, A. (1977). Relationships between oral reading rates for letters, words, and simple text in the development of reading achievement. *Reading Research Quarterly*, *13*(2), 223–253. Available from http://www.jstor.org/stable/747307

Braze, D., Tabor, W., Shankweiler, D., & Mencl, W. (2007). Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities*, *40*(3), 226.

Broadbent, D. (1967). Word-frequency effect and response bias. *Psychological Review*, *74*(1), 1–15. Available from http://www.sciencedirect.com/science/article/B6X04-4NR7FDN-1/2/b7c013c23e8df6273e930e0f61c147ad

Broadbent, D., & Broadbent, M. (1975). Some further data concerning the word frequency effect. *Journal of Experimental Psychology: General*, *104*(4), 297–308. Available from http://www.sciencedirect.com/science/article/B6X07-4NVH9F4-1/2/d5a1fb82653866ae3c217ea5043df00d

Broadbent, D., & Gregory, M. (1968). Visual perception of words differing in letter digram frequency. *Journal of Verbal Learning and Verbal Behavior*, *7*(2), 569–571. Available from http://www.sciencedirect.com/science/article/B7MD4-4H3SDHF-1T/2/2ea26079ee91ed16c1ba05ace45ce3c4

Brown, J. I., Bennett, F. M., & Hanna, G. (1981). *The Nelson Denny reading test (form E)*. Chicago: Riverside Publishing.

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology*, *97*(3), 165–182. Available from http://www.sciencedirect.com/science/article/B6WJ9-4NDDT8K-1/2/fbdf32744abadf616004e0dcd777127c

Connine, C., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6), 1084–1096.

Cunningham, A., & Stanovich, K. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology*, *83*(2), 264–274.

Cunningham, A., & Stanovich, K. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–945.

Cunningham, T., Healy, A., Kanengiser, N., Chizzick, L., & Willitts, R. (1988). Investigating the boundaries of reading units across ages and reading levels. *Journal of Experimental Child Psychology*, *45*(2), 175–208. Available from http://www.sciencedirect.com/science/article/B6WJ9-4D6YWVR-J/2/6c02f7297f33b134242ca46f70455f0e

Dahan, D., Magnuson, J., & Tanenhaus, M. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive psychology*, *42*(4), 317–367.

Feng, G. (2006). Eye movements as time-series random variables: A stochastic model of eye movement control in reading. *Cognitive Systems Research*, *7*(1), 70–95. Available from http://www.sciencedirect.com/science/article/B6W6C-4HB4DN2-1/2/9147d9856f32efe160495602da9db50c

Finucci, J. M., Isaacs, S. D., Whitehouse, C. C., & Childs, B. (1982). Empirical validation of reading and spelling quotients. *Developmental Medicine & Child Neurology*, *24*(s45), 733–744. Available from http://dx.doi.org/10.1111/j.1469-8749.1982.tb13695.x

Finucci, J. M., Whitehouse, C. C., Isaacs, S. D., & Childs, B. (1984). Derivation and validation of a quantitative definition of specific reading disability for adults. *Developmental Medicine & Child Neurology*, *26*(2), 143–153. Available from http://dx.doi.org/10.1111/j.1469-8749.1984.tb04425.x

Forster, K., & Chambers, S. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*(6), 627–635.

Francis, W., & Kucera, H. (1982). *Frequency analysis of English useage: lexicon and grammar*. Boston: Houghton Mifflin.

Frederiksen, J., & Kroll, J. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(3), 361–379. Available from http://www.sciencedirect.com/science/article/B6X08-4NRKDCC-7/2/8054d8e23824a6959131e3d0035d8243

Fuchs, L., Fuchs, D., Hosp, M., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239–256. Available from http://www.informaworld.com/10.1207/S1532799XSSR0503_3

Gernsbacher, M. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*(2), 256–281. Available from http://www.sciencedirect.com/science/article/B6X07-4NRKDD7-7/2/50f0aed9e406a6212a94c45d0cf4360b

Greenberg, S. N., Healy, A., Koriat, A., & Kreiner, H. (2004). The go model: A reconsideration of the role of structural units in guiding and organizing text on line. *Psychonomic* *bulletin & review*, *11*(3), 428-433 N1 –. Available from http://pbr.psychonomic-journals.org/content/11/3/428.abstract

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23. Available from http://www.sciencedirect.com/science/article/B6VH9-4HNYMV0-1/2/e2617c8b4b70ecc26e81478a269f49fa%20AB

Hart, L. (2005). *A training study using an artificial orthography: Effects of reading experience, lexical quality, and text comprehension in L1 and L2*. Phd thesis, University of Pittsburgh.

Healy, A. (1976). Detection errors on the word the : Evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(2), 235–242. Available from http://www.sciencedirect.com/science/article/B6X08-4NRKDBW-9/2/54c4cf09bf8b4ba4d6daae050d06fc77

Healy, A. (1980). Proofreading errors on the word the: New evidence on reading units. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(1), 45–57. Available from http://www.sciencedirect.com/science/article/B6X08-4NV6PPT-4/2/c133f9c291ca5bfed7fb9bfb420af3bc

Healy, A. (1981). Cognitive processes in reading text. *Cognition*, *10*(1-3), 119–126.

Howes, D., & Solomon, R. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology*, *41*(6), 401–410. Available from http://www.sciencedirect.com/science/article/B8JB9-4NTJPHW-2/2/5379a4550f5abe49120b572311166e8f

Jenkins, J., Stein, M., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal*, *21*(4), 767.

Juel, C. (1980). Comparison of word identification strategies with varying context, word type, and reader skill. *Reading Research Quarterly*, *15*(3), 358–376. Available from http://www.jstor.org/stable/747420

Juhasz, B., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1312–1318.

Kacinik, N., Shears, C., & Chiarello, C. (2000). Familiarity for nouns and verbs: not the same as, and better than, frequency. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (p. 1035). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kaiser, H. (1958, Sep). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 187–200. Available from http://dx.doi.org/10.1007/BF02289233

Kaiser, H. (1974, Mar). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36. Available from http://dx.doi.org/10.1007/BF02291575

Katz, L., Lee, C., Tabor, W., Frost, S., Mencl, W., Sandak, R., et al. (2005). Behavioral and neurobiological effects of printed word repetition in lexical decision and naming. *Neuropsychologia*, *43*(14), 2068–2083. Available from http://www.sciencedirect.com/science/article/B6T0D-4G0YTS1-3/2/d030cb4db56e7fa610a1e8d5771a805d

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1), 262 – 284. Available from http://www.informaworld.com/10.1080/09541440340000213

LaBerge, D., & Samuels, S. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, *6*(2), 293–323.

Landi, N. (2005). *Behavioral and electrophysiological investigations of semantic processing in skilled and less-skilled comprehenders*. Phd thesis, University of Pittsburgh.

Lee, H., Rayner, K., & Pollatsek, A. (1999). The time course of phonological, semantic, and orthographic coding in reading: Evidence from the fast priming technique. *Psychometric Bulletin & Review*, *5*, 624–634.

Lefly, D., & Pennington, B. (2000). Reliability and validity of the adult reading history questionnaire. *Journal of Learning Disabilities*, *33*(3), 286–296. Available from http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=3109412&site=ehost-live

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71–102.

McDonald, S., Carpenter, R., & Shillcock, R. (2005). An anatomically constrained, stochastic model of eye movement control in reading. *Psychological Review*, *112*(4), 814–840. Available from http://www.sciencedirect.com/science/article/B6X04-4HKMCCM-4/2/d475627954acf92ad0d7e9b82c8fe5dd

McKeown, M., Beck, I., Omanson, R., & Perfetti, C. A. (1983). The effects of long-term vocabulary instruction on reading comprehension: A replication. *Journal of Reading Behavior*, *15*(1), 3–18. Available from http://www.informaworld.com/10.1080/10862968309547474

McKeown, M., Beck, I., Omanson, R., & Pople, M. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly*, *20*(5), 522–535. Available from http://www.jstor.org/stable/747940

Monsell, S., Doyle, M., & Haggard, P. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, *118*(1), 43–71.

Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly*, *20*(2), 233–253. Available from http://www.jstor.org/stable/747758

Nathan, R., & Stanovich, K. (1991). The causes and consequences of differences in reading fluency. *Theory into Practice*, *30*(3), 176–184. Available from http://www.jstor.org/stable/1476879

Nelson, J. R., Balass, M., & Perfetti, C. A. (2005, Jun). Differences between written and spoken input in learning new words. *Written Language and Literacy*, *8*(2), 25–44.

Nelson, N. J., & Denny, E. C. (1973). *Nelson Denny Reading Test.* Boston: Houghton Mifflin Co.

Oldfield, R., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*(4), 273–281. Available from http://www.informaworld.com/10.1080/17470216508416445

Olson, R., Wise, B., Conners, F., Rack, J., & Fulker, D. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. *Journal of Learning Disabilities*, *22*(6). Available from http://search.ebscohost.com/

login.aspx?direct=true&db=aph&AN=4739719&site=ehost-live

Perfetti, C. A. (2007a). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. Available from http://www.informaworld.com/10.1080/10888430701530730

Perfetti, C. A. (2007b, Oct). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *4*, 357–383.

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and masked priming. *Journal of Memory and Language*, *30*, 473–485.

Perfetti, C. A., Bell, L., & Delaney, S. (1988). Automatic phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language*, *27*, 59–70.

Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. Gorfien (Ed.), *On the consequences of meaning selection* (p. 67-86). Washington, D.C.: American Psychological Association.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy.* Amsterdam: John Benjamins.

Perfetti, C. A., Wlotko, E., & Hart, L. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1281–1292. Available from http://www.sciencedirect.com/science/article/B6X09-4HYWS5J-B/2/d870147158eddb5e9137dc8fbff1bfb6

Pinnell, G. S., Pikulski, J. J., Wixson, K. K., Campbell, J. R., Gough, P. B., & Beatty, A. S. (1995). *Listening to children read aloud.* Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Pollatsek, A., Perea, M., & Binder, K. S. (1999, Aug). The effects of "neighborhood size" in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 1142–58.

Posner, M., DiGirolamo, G., & Fernandez-Duque, D. (1997). Brain mechanisms of cognitive skills. *Consciousness and Cognition*, *6*(2-3), 267–290. Available from http://www.sciencedirect.com/science/article/B6WD0-45MFWFN-N/2/98803ff861bf5e900b8474728ead1daa%20AB

Rasinski, T., Padak, N., McKeon, C., Wilfong, L., Friedauer, J., & Heim, P. (2005). Is reading fluency a key for successful high school reading? *Journal of Adolescent & Adult Literacy*, *49*(1), 22–27. Available from http://www.jstor.org/stable/40009266

Raven, J. C. (1960). *Guide to the Standard Progressive Matrices*. London: H K Lewis.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, *41*(2), 211–236.

Rayner, K., Sereno, S., Lesch, M., & Pollatsek, A. (1995). Phonological codes are automatically activated during reading: Evidence from an eye movement priming paradigm. *Psychological Science*, *6*(1), 26–32. Available from http://www.jstor.org/stable/40062872

Rayner, K., Sereno, S., & Raney, G. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(5), 1188–1200.

Reichle, E., & Perfetti, C. A. (2003). Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading*, *7*(3), 219–237. Available from http://www.informaworld.com/10.1207/S1532799XSSR0703_2

Reichle, E., Pollatsek, A., & Rayner, K. (2006, Jan). Ez reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*. Available from http://linkinghub.elsevier.com/retrieve/pii/S1389041705000239

Reilly, R., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research*, *7*(1), 34–55.

Rice, G. A., & Robinson, D. O. (1975). The role of bigram frequency in the perception of words and nonwords. *Memory & Cognition*, *3*, 513–518.

Richter, E., Engbert, R., & Kliegl, R. (2006). Current advances in swift. *Cognitive Systems Research*, *7*(1), 23–33. Available from http://www.sciencedirect.com/science/article/B6W6C-4H21NHS-2/2/ecaec8c45173c7a26df2752e316221c5

Samuels, S. J. (2002). Reading Fluency: Its Development and Assessment. In A. E. Farstrup & S. J. Samuels (Eds.), *What Research Has to Say About Reading Instruction* (Third ed., pp. 166–184). Newark, DE: International Reading Association.

Samuels, S. J., LaBerge, D., & Bremer, C. (1978). Units of word recognition: Evidence for developmental changes. *Journal of Verbal Learning and Verbal Behavior*, *17*(6), 715–720.

Samuels, S. J., Miller, N., & Eisenberg, P. (1979). Practice effects on the unit of word recognition. *Journal of Educational Psychology*, *71*(4), 514–520. Available from http://www.sciencedirect.com/science/article/B6WYD-4NV6XDS-H/2/b7c2a292b1bb2f274a67fc89ffaaf9c0

Sandak, R., Mencl, W., Frost, S., & Pugh, K. (2004). The neurobiological basis of skilled and impaired reading: Recent findings and new directions. *The Cognitive Neuroscience of Reading: A Special Issue of Scientific Studies of Reading*, *8*(3), 273–292.

Schilling, H. H., Rayner, K., & Chumbley, J. I. (1998, Nov). Comparing naming, lexical decision, and eye fixation times: word frequency effects and individual differences. *Memory & Cognition*, *26*(6), 1270–81.

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Zoning out while reading: Evidence for dissociations between experience and metaconsciousness. In D. T. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (p. 204-226). Cambridge: MIT Press.

Sears, C., Hino, Y., & Lupker, S. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 876–900.

Share, D. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218. Available from http://www.sciencedirect.com/science/article/B6T24-3YYTDVX-5/2/36b036e775a8dcc7cddbdf4d459f1d91

Solomon, R., & Postman, L. (1952). Frequency of usage as a determinant of recognition thresholds for words. *Journal of Experimental Psychology*, *43*(3), 195–201. Available from http://www.sciencedirect.com/science/article/B8JB9-4P4H8SF-5/2/87fa5c0e7356458a92c4a3ba34451da3

Spieler, D., & Balota, D. A. (2000). Factors influencing word naming in younger and older adults. *Psychology and Aging*, *15*(2), 225–231.

Stahl, S., & Fairbanks, M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of educational research*, *56*(1), 72.

Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407. Available from http://www.jstor.org/stable/747612

Stanovich, K., & Cunningham, A. (1992, Nov). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognitition*, *20*(1), 51–68.

Stanovich, K., & West, R. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*(4), 402–433. Available from http://www.jstor.org/stable/747605

Stanovich, K., West, R., & Harrison, M. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Developmental Psychology*, *31*(5), 811–826. Available from http://www.sciencedirect.com/science/article/B6WYC-46RVNGS-9/2/a8337f35685ad7c84088cdbefba273b3

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. *The Oxford handbook of psycholinguistics*, 327–342.

Verhoeven, L., & Leeuwe, J. van. (2008). Prediction of the development of reading comprehension: a longitudinal study. *Applied Cognitive Psychology*, *22*(3), 407–423. Available from http://dx.doi.org/10.1002/acp.1414

West, R., & Stanovich, K. (1991). The incidental acquisition of information from reading. *Psychological Science*, *2*(5), 325.

West, R., Stanovich, K., & Mitchell, H. (1993). Reading in the real world and its correlates. *Reading Research Quarterly*, *28*(1), 35–50. Available from http://www.jstor.org/stable/747815

Whaley, C. (1978). Word–nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, *17*(2), 143–154.

White, S. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 205–223.

Williams, C. C., Perea, M., Pollatsek, A., & Rayner, K. (2006, Aug). Previewing the neighborhood: the role of orthographic neighbors as parafoveal previews in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 1072–82.

Williams, R., & Morris, R. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology*, *16*(1), 312–339. Available from http://www.informaworld.com/10.1080/09541440340000196

Winnick, W., & Kressel, K. (1965). Tachistoscopic recognition thresholds, paired-associate learning, and free recall as a function of abstractness-concreteness and word frequency. *Journal of Experimental Psychology*, *70*(2), 163–168. Available from http://www.sciencedirect.com/science/article/B8JB9-4NT9X2S-6/2/c4ef236b5c6a881f998920b1488b8467

Yates, M., Friend, J., & Ploetz, D. (2008). The effect of phonological neighborhood density on eye movements during reading. *Cognition*, *107*(2), 685–692. Available from http://www.sciencedirect.com/science/article/B6T24-4PMJB1P-1/2/e88f817879a6f5a74f9a4125d8d10a95